TOPICS IN
NUMBER THEORY

VOLUME II

This book is in the

ADDISON-WESLEY SERIES IN MATHEMATICS

Eric Reissner, *Consulting Editor*

TOPICS IN NUMBER THEORY

VOLUME II

by

WILLIAM JUDSON LeVEQUE

Department of Mathematics
University of Michigan

PREFACE

This book is a treatment of some advanced topics in the theory of numbers. It was written to follow the author's *"Topics in Number Theory, Volume I,"* in which elementary number theory is presented.

The level of mathematical maturity required for Volume II is much higher than for Volume I. Moreover, results obtained in Volume I are used freely, and in several of the chapters a knowledge of specific topics in various other branches of mathematics is assumed. In particular, knowledge of the theory of symmetric polynomials, as well as the rule for multiplying determinants, is needed for the algebraic theory in Chapter 3, and the theory of analytic functions is used both in the theorem of Schneider in Chapter 5 and in the investigation of the distribution of primes in Chapter 7. There seemed to be no point in assuming background unnecessarily, however, so I have included brief discussions of groups and matrices, on a very elementary level, in Chapter 1.

The treatment of quadratic forms, admittedly shallow, has been based on the properties of the modular group for two reasons. In the first place, the geometric interpretation makes the usual definition of reduced forms seem quite natural, while no real insight is afforded by merely listing an unmotivated set of inequalities. In the second place, this treatment provides a simple illustration of the power of the theory associated with elliptic functions, which is of considerable importance in modern number theory. Such methods are not often taught in American universities, and I hope that this treatment may serve to stimulate interest in them.

To the best of my knowledge, the algebraic form of the Thue-Siegel-Roth theorem given in Chapter 4 has not previously appeared in print.

W. J. L.

Ann Arbor, Michigan
November 1955

CONTENTS

CHAPTER 1

BINARY QUADRATIC FORMS

1-1 Introduction. One of the subjects treated in elementary number theory is the possibility of representing a positive integer as a sum of two squares.* The expression $x^2 + y^2$ which is of interest for this problem is a special case of the general *binary quadratic form*

$$f(x, y) = ax^2 + bxy + cy^2. \tag{1}$$

(This in turn is a special case of the n-ary m-ic form, which is a homogeneous polynomial of degree m in n variables.) Systematic research in quadratic forms was begun by Gauss, and has since been extensively pursued. We shall not go very deeply into the subject, but prefer instead to develop general methods whose usefulness is not limited to the theory of quadratic forms, nor even to the theory of numbers.

Suppose that in (1) we make the linear homogeneous substitution

$$\begin{aligned} x &= \alpha x' + \beta y', \\ y &= \gamma x' + \delta y', \end{aligned} \tag{2}$$

where α, β, γ, and δ are integers and $D = \alpha\delta - \beta\gamma \neq 0$. Solving for x' and y' gives

$$x' = \frac{\delta}{D} x - \frac{\beta}{D} y,$$

$$y' = -\frac{\gamma}{D} x + \frac{\alpha}{D} y, \tag{3}$$

so it is only in case $D = \pm 1$ that to each integer pair x, y corresponds an integer pair x', y' and conversely. We shall eventually suppose

---

*See, for example, LeVeque, *Topics in Number Theory*, vol. I, (Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1956), Chapter 7. So much use will be made of the results obtained in this book that it will be referred to henceforth simply as Volume I.

1

that $D = +1$, for reasons that will appear later. Then

$$x' = \delta x - \beta y,$$
$$y' = -\gamma x + \alpha y. \tag{4}$$

Substituting (2) into (1), we have a new form in x' and y',

$$g(x', y') = Ax'^2 + Bx'y' + Cy'^2,$$

where

$$A = a\alpha^2 + b\alpha\gamma + c\gamma^2,$$
$$B = 2a\alpha\beta + b(\alpha\delta + \beta\gamma) + 2c\gamma\delta, \tag{5}$$
$$C = a\beta^2 + b\beta\delta + c\delta^2.$$

If for suitable integral values of x and y we have $f(x, y) = n$, then, for the corresponding values of x' and y' determined by equation (4), $g(x', y') = n$.

It thus appears that, as far as questions of representation are concerned, it would be senseless duplication to consider $f(x, y)$ and $g(x', y')$ separately; every integer represented by f is also represented by g, and conversely. This leads us to call f and g *equivalent*, and to write $f \sim g$, if one can be obtained from the other by a unimodular linear substitution with integral coefficients,

$$x = \alpha x' + \beta y', \qquad \alpha\delta - \beta\gamma = 1. \tag{6}$$
$$y = \gamma x' + \delta y',$$

This in turn brings up one of the principal problems of this chapter: how to decide whether two given forms are equivalent.

Substitution (2) is described quite adequately by specifying coefficients α, β, γ, and δ; that is, by writing the *matrix*

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}.$$

That does not represent a number, of course; it is simply a the coefficients of the substitution, in the order in which they occur in (2). However, we can give names to these matrices, and deduce certain of their properties from the corresponding properties of the associated substitutions. Thus if

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{and} \quad M' = \begin{pmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{pmatrix},$$

then we shall say that M and M' are equal if and only if they correspond to the same substitution, that is,

$$\alpha = \alpha', \qquad \beta = \beta', \qquad \gamma = \gamma', \qquad \delta = \delta'.$$

If for arbitrary M and M' we apply the corresponding substitutions successively, so that

$$x = \alpha x' + \beta y', \qquad x' = \alpha' x'' + \beta' y'',$$

$$y = \gamma x' + \delta y', \qquad y' = \gamma' x'' + \delta' y'',$$

we could accomplish the same thing by the single substitution

$$x = (\alpha\alpha' + \beta\gamma')x'' + (\alpha\beta' + \beta\delta')y'',$$

$$y = (\gamma\alpha' + \delta\gamma')x'' + (\gamma\beta' + \delta\delta')y''.$$

Thus, if by the *product* MM' of two matrices we mean the matrix of this latter substitution, we must define

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \alpha' & \beta' \\ \gamma' & \delta' \end{pmatrix} = \begin{pmatrix} \alpha\alpha' + \beta\gamma' & \alpha\beta' + \beta\delta' \\ \gamma\alpha' + \delta\gamma' & \gamma\beta' + \delta\delta' \end{pmatrix}.$$

Thus the product has as element in the ith row and jth column, for each i and j, the sum of the products of the elements of the ith row of the first matrix with the corresponding elements of the jth column of the second matrix. Moreover, if the *determinant* of a matrix is defined as

$$\det \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{vmatrix} \alpha & \beta \\ \gamma & \delta \end{vmatrix} = \alpha\delta - \beta\gamma,$$

it requires only a routine calculation to show that

$$\det M \cdot \det M' = \det (MM').$$

It is to be noticed that, in general, $MM' \neq M'M$, although

$$M(M'M'') = (MM')M''.$$

Since the substitutions given by (2) and (3) are invers[e to each] other, it is natural to call the matrix of (3) the *inverse* of the m[atrix] of (2), and to designate it by M^{-1}. Then $MM^{-1} = M^{-1}M = I$, where

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{7}$$

I is called the *identity matrix;* it corresponds to the trivial substitu-

tion $x = x'$, $y = y'$, and has the property that $MI = IM = M$ for every M. A square matrix has an inverse if and only if its determinant is different from zero.

Finally, we designate by $\overline{M}$ the *transpose* of M, obtained by interchanging rows and columns in M:

$$\text{if } M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \text{then } \overline{M} = \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix}.$$

The transpose of a product is the product of the transposes, in reverse order:

$$(\overline{MM'}) = \overline{M}'\overline{M}.$$

Also, the transpose of the inverse is equal to the inverse of the transpose:

$$(\overline{M})^{-1} = \overline{M^{-1}}.$$

Matrices need not be square. Thus

$$\text{if } X = (x\ y), \quad \text{then } \overline{X} = \begin{pmatrix} x \\ y \end{pmatrix};$$

note, however, that nonsquare matrices have neither determinants nor inverses.

The importance of this algebra of matrices to our present purpose lies in the fact that if

$$F = \begin{pmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{pmatrix},$$

then

$$XF\overline{X} = (x\ y) \begin{pmatrix} a & \frac{1}{2}b \\ \frac{1}{2}b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = (ax + \tfrac{1}{2}by \quad \tfrac{1}{2}bx + cy)\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= (ax^2 + bxy + cy^2).$$

Although it is a slight abuse of language, it is convenient and in the present context harmless to identify a one-by-one matrix with the element itself, so we write

$$f(x, y) = XF\overline{X}.$$

F is called the *matrix* of the form, and $\Delta = 4 \cdot \det F = 4ac - b^2$ is called the *discriminant* of the form.

In terms of matrices, the substitution equations (2) and (3) can be written as

$$X = X'\overline{M} \quad \text{and} \quad X' = X\overline{M}^{-1},$$

respectively. Thus

$$f(x, y) = XF\overline{X} = (X'\overline{M})F(\overline{X'\overline{M}}) = (X'\overline{M})F(M\overline{X}')$$
$$= X'(\overline{M}FM)\overline{X}',$$

so that the matrix of g is $G = \overline{M}FM$. (The reader might test his ability to manipulate matrices by showing that the last equation is in agreement with equations (5)). Multiplying both sides of the equation $G = \overline{M}FM$ by $\overline{M}^{-1}$ on the left and M^{-1} on the right, we have

$$\overline{M}^{-1}GM^{-1} = \overline{M}^{-1}(\overline{M}FM)M^{-1} = (\overline{M}^{-1}\overline{M})F(MM^{-1}) = F.$$

If $\det M = 1$, then also $\det \overline{M} = 1$, and

$$\det G = \det (\overline{M}FM) = \det \overline{M} \cdot \det F \cdot \det M = \det F,$$

so that the discriminant of a form is not changed by a unimodular substitution.

In summary, a form with matrix F is equivalent to a form with matrix G if and only if there is a matrix M such that $G = \overline{M}FM$ and $\det M = \mp 1$; equivalent forms have the same discriminant and represent the same integers.

The relation of "equivalence," as used here, is an equivalence relation in the technical sense.* For it is clear that

(a) $f \sim f$: $F = \overline{I}FI$;

(b) $f \sim g$ implies $g \sim f$: $G = \overline{M}FM$ implies $F = \overline{M}^{-1}GM^{-1}$;

(c) $f \sim g$ and $g \sim h$ implies $f \sim h$: $G = \overline{M}FM$ and $H = \overline{M}'GM'$

implies $H = \overline{M}'\overline{M}FMM' = (\overline{MM'})F(MM')$.

Thus all the forms equivalent to a given one are equivalent to each other, and the set of all forms splits up into equivalence classes, any two elements in one class being equivalent, and elements from different classes being inequivalent. (The equivalence classes for the relation of congruence (mod m) are simply the residue classes modulo m.) Just as we chose a system of representatives of the various residue classes modulo m, we would like to pick a system of representative forms, one from each class. It is the object of the next two sections to develop machinery by which such *reduced* forms can be obtained in a natural way.

---

*See, for example, Volume I, Section 3–1.

primitive root of q is a generator of $M(q)$; thus Theorem 4–11 is a statement of the fact that $M(q)$ is cyclic (consists of the powers of a single element) if and only if $q = 1, 2, 4, p^{\alpha}, 2p^{\alpha}$. If $\alpha > 2$, $M(2^{\alpha})$ has two generators, -1 and 5; for example, modulo 16, the powers of 5 are 5, 9, 13, and 1, and these numbers, together with their negatives, form a reduced residue system.

At present we shall do no more with finite groups, but turn our attention instead to the much more complicated multiplicative group of all two-by-two matrices with integral entries and unit determinants. This infinite group, which will be designated here by Γ, is called the *modular group*. To show that Γ is a group, we verify properties (a) through (d) above. The system is obviously closed under multiplication, since the determinant of a product is the product of the determinants of the factors. The associative property has already been verified. The identity element of Γ is I, as defined in (7). The inverse of any element

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

is

$$\begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix},$$

since

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix} = \begin{pmatrix} \delta & -\beta \\ -\gamma & \alpha \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = I.$$

The group Γ differs from the other examples mentioned in that it is noncommutative, since in general $MM' \neq M'M$. (Abstractly, G is said to be a *commutative* or *abelian* group if $a \circ b = b \circ a$ for every a and b in G.)

1–3 The modular group. The properties of Γ could all be developed by the use of algebra alone; we prefer instead to build up the theory with the help of a simple geometric interpretation. It is now convenient to reverse the roles of the accented and unaccented variables in the equations (2); this new notation will be used throughout the discussion of the modular group, but the original system will be reverted to when quadratic forms are again considered. To keep matters straight, (2) will be termed a *substitution*, while the modified equations will be called a *transformation*. Putting $z = x/y$ and

$z' = x'/y'$, we get

$$z' = \frac{\alpha z + \beta}{\gamma z + \delta}. \tag{8}$$

So far nothing essential has been accomplished. The crucial point lies in allowing z to range over all complex numbers, rather than the real rationals to which it was formerly restricted. Then equation (8) can be regarded as defining a transformation or mapping of the complex z-plane into the z'-plane. Somewhat more than this can be said: if

$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

is in Γ, so that $\det M = 1$, a simple calculation shows that the imaginary parts of z and z' have the same sign. In other words, (8) maps the upper half of the z-plane (i.e., the region where the imaginary part of z is positive) into the upper half of the z'-plane, and the lower half into the lower half. Hereafter, we restrict attention to the upper half planes.

It is convenient to identify the z- and z'-planes, and to think of (8) as sending each point z of the upper half U of the complex plane into another point z' of U. We also identify the elements of Γ with the corresponding transformations (8), which has the effect of identifying the matrices

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -\alpha & -\beta \\ -\gamma & -\delta \end{pmatrix}.$$

In accordance with the earlier definition of equivalence, two points z and z' of U will be called *equivalent* if one can be mapped into the other by a transformation of Γ. As usual, this assigns each point of U to an equivalence class; two elements of the same class are equivalent, and elements from different classes are inequivalent. A region R of U is called a *fundamental region* if no two of its points are equivalent, while every point of U is equivalent to a point of R; in other words, R constitutes a complete system of representatives of the above equivalence classes. It would be more precise to refer to R as a *fundamental region of the group* Γ, since two points may be equivalent with respect to one group of transformations but not with respect to another. For example, it is clear that a fundamental region R' of a subgroup Γ' of Γ contains a fundamental region of Γ itself, if both

regions exist. For any point in U, being equivalent to some point of R' under the transformations of Γ', is *a fortiori* equivalent to the same point of R' under the transformations of the larger group Γ. It may not be true, however, that any two points of R' are inequivalent with respect to Γ.

THEOREM 1–1. *The region R in U composed of all points z such that $-\frac{1}{2} \leq \operatorname{Re} z < \frac{1}{2}$ and either $|z| > 1$, or else $|z| = 1$ and $-\frac{1}{2} \leq \operatorname{Re} z \leq 0$, is a fundamental region of Γ.* (See Fig. 1–1.)

Proof: First note that Γ has the subgroup Γ_0 of all integral translations $z' = z + \beta$. For the associated matrix

$$\begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix}$$

has determinant 1, the identity transformation $z' = z$ is in Γ_0, the inverse of $z' = z + \beta$ is $z' = z - \beta$ and is in Γ_0, and the result of making two translations is again a translation. Γ_0 is cyclic, being generated by

$$z' = z + 1. \tag{9}$$
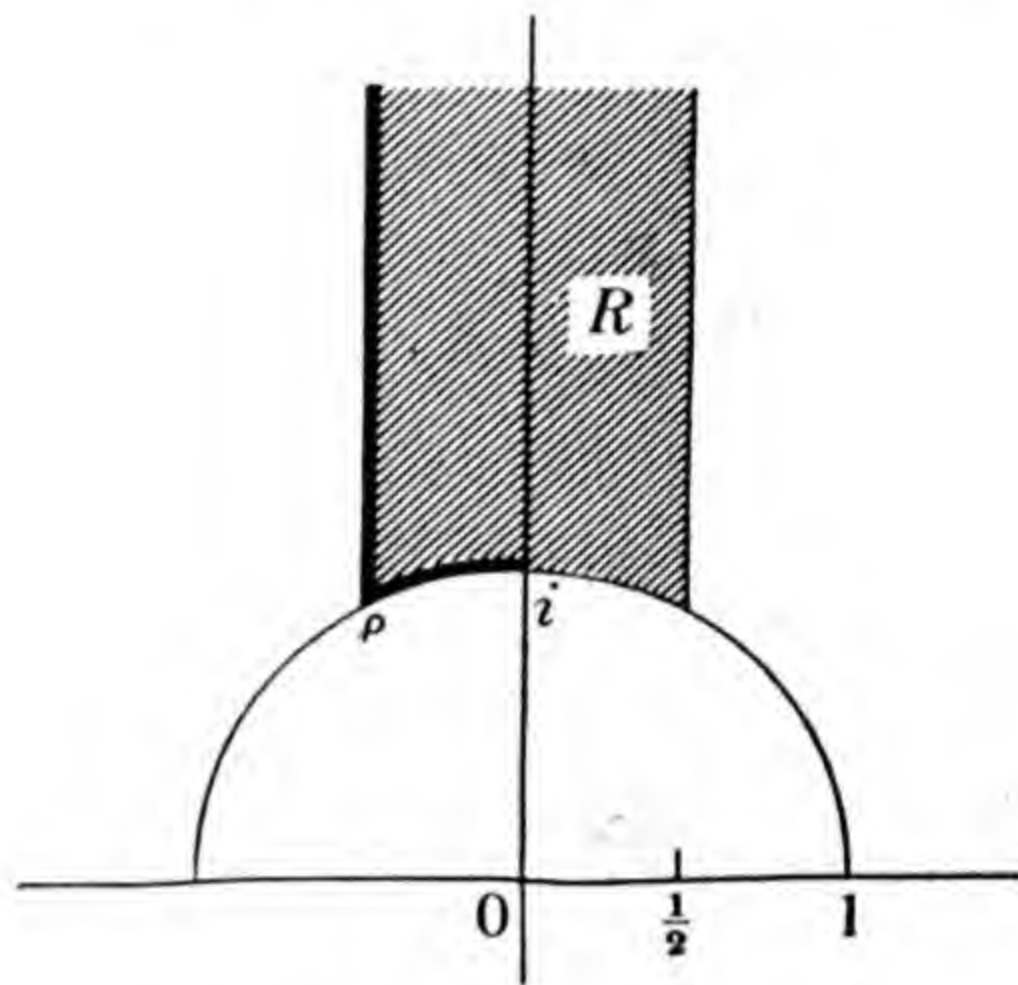
FIGURE 1–1

As a fundamental region of Γ_0 we could choose any infinite strip in U of unit width, extending parallel to the imaginary axis from the real axis. We take the following one:

$$R_0: \quad \operatorname{Im} z > 0, \quad -\frac{1}{2} \leq \operatorname{Re} z < \frac{1}{2}.$$

From the remark preceding the theorem, R_0 must contain a fundamental region of Γ if any exists. R_0 is not itself a fundamental region of Γ, however, for the point $i/2$ of R_0 is transformed into the point $2i$ of R_0 by the transformation

$$z' = -\frac{1}{z}. \tag{10}$$

With each transformation

$$T: \quad z' = \frac{\alpha z + \beta}{\gamma z + \delta}$$

with $\gamma \neq 0$, there is associated the circle $C(T): |\gamma z + \delta| = 1$, with center at $-\delta/\gamma$ and radius $1/|\gamma| \leq 1$. Now

$$\gamma z' - \alpha = \gamma \frac{\alpha z + \beta}{\gamma z + \delta} - \alpha = \frac{\beta\gamma - \alpha\delta}{\gamma z + \delta} = -\frac{1}{\gamma z + \delta},$$

so that $C(T)$ is transformed by T into $|\gamma z - \alpha| = 1$, which, by (3), is $C(T^{-1})$. More importantly, the exterior of $C(T)$ goes into the interior of $C(T^{-1})$. It is simple to deduce from this that no two points of the region R described in the theorem are equivalent. Certainly no point of R is mapped into another by an element of Γ_0. But if T is not in Γ_0, then $\gamma \neq 0$, and since the interior of R is external to all the circles $C(T)$ (inasmuch as they all have radii ≤ 1 and are centered at real points), any interior point of R is mapped by T into an interior point of one of these circles, and hence into a point outside R.

The arc $A: |z| = 1$, $-\frac{1}{2} \leq \mathrm{Re}\, z \leq 0$, which forms part of the boundary of R, is also completely exterior to all the circles $C(T)$ except $|z| = 1$ and $|z + 1| = 1$. The circle $|z| = 1$ is associated with transformations

$$z' = \frac{\alpha z + \beta}{z},$$

and since the determinant must be 1, $\beta = -1$ and

$$z' = \frac{\alpha z - 1}{z} = \alpha - \frac{1}{z}, \qquad |z' - \alpha| = \frac{1}{|z|}.$$

If z is a point of A, $|z' - \alpha| = 1$, and so z' is not in R unless $\alpha = 0$ or -1. If $\alpha = 0$ we have the transformation (10), which sends A onto the arc $|z| = 1, 0 \leq \mathrm{Re}\, z \leq \frac{1}{2}$; this arc has only the point i in common with R, and i goes into itself. (This means that i is equivalent to itself in two different ways: $z' = z$ and $z' = -1/z$.) If $\alpha = -1$, A goes into the arc $|z + 1| = 1$, $-1 \leq \mathrm{Re}\, z \leq -\frac{1}{2}$; these two arcs have just $\rho = -\frac{1}{2} + i\sqrt{3}/2$ in common, and ρ goes into itself.

The circle $|z + 1| = 1$ is associated with transformations

$$z' = \frac{\alpha z + \beta}{z + 1} = \frac{\alpha z + (\alpha - 1)}{z + 1} = \alpha - \frac{1}{z + 1}.$$

If $|z| = 1$, then

$$\left| \frac{z' - (\alpha - 1)}{z' - \alpha} \right| = 1,$$

$$|z' - (\alpha - 1)| = |z' - \alpha|,$$

$$\operatorname{Re} z' = \alpha - \tfrac{1}{2},$$

and z' is not in R unless $\alpha = 0$. Under the transformation

$$z' = -\frac{1}{z + 1},$$

the arc A goes into the line segment $\operatorname{Re} z = -\tfrac{1}{2}$, $\tfrac{1}{2} \leq \operatorname{Im} z \leq \sqrt{3}/2$; the arc and the segment have just ρ in common, and ρ goes into itself. We have thus shown that no two points of R are equivalent, and have incidentally obtained the following result, which will be useful later.

THEOREM 1–2. *The point $\rho = (-1 + i\sqrt{3})/2$ is mapped into itself by the three transformations*

$$z' = z, \qquad z' = -\frac{1}{1 + z}, \qquad \text{and} \qquad z' = -\frac{z + 1}{z},$$

and by no others. The point i is mapped into itself by the two transformations

$$z' = z \qquad \text{and} \qquad z' = -\frac{1}{z},$$

and by no others. Any point of R different from ρ and i is mapped into itself only by the identity transformation $z' = z$.

To complete the proof of Theorem 1–1, we must show that any point z in U is the image of a point in R under a transformation of Γ. We do this by finding a finite sequence of transformations such that if they are successively applied to z, the final point z' is in R. Then the inverse of the product of these transformations maps z' back into z.

Designate by S the generator (9) of Γ_0, and by W the transformation (10). Let z be a point of U not in R. Then for some integer n_1, which may be positive, negative, or zero, $z_1 = S^{n_1}z = z + n_1$ is in R_0, the fundamental domain of Γ_0. If z_1 is in R, we are finished. If $|z_1| = 1$ but $0 < \operatorname{Re} z \leq \tfrac{1}{2}$, then Wz_1 is in R. If $|z_1| < 1$, then $z_2 = Wz_1$ has modulus greater than 1. In fact, if $z_1 = x_1 + iy_1$,

$$z_2 = \frac{-x_1 + iy_1}{x_1{}^2 + y_1{}^2} = x_2 + iy_2, \qquad -\tfrac{1}{2} \le x_1 < \tfrac{1}{2},$$

so that $\operatorname{Im} z_2 > \operatorname{Im} z_1$, and if $y_1 \le \tfrac{1}{2}$, then $\operatorname{Im} z_2 \ge 2 \operatorname{Im} z_1$, since then $x_1{}^2 + y_1{}^2 \le \tfrac{1}{2}$. If z_2 is in R, we are through. If not, there is a suitable exponent n_2 such that $z_3 = S^{n_2} z_2$ is in R_0, and $\operatorname{Im} z_3 = \operatorname{Im} z_2$. If z_3 is not in R, we can apply W again, and get $z_4 = W S^{n_2} W S^{n_1} z$. What we must show is that after finitely many steps, this process leads to a point in R.

As long as $y_k \le \tfrac{1}{2}$, we will have $y_{k+1} \ge 2 y_k$, if $z_{k+1} = W z_k$. Starting with a positive number (the imaginary part of z), a finite number of doublings will produce a number larger than $\tfrac{1}{2}$. So suppose that we have obtained a $z_k = x_k + i y_k$ such that

$$-\tfrac{1}{2} \le x_k < \tfrac{1}{2}, \qquad y_k > \tfrac{1}{2}, \qquad x_k{}^2 + y_k{}^2 < 1. \tag{11}$$

Then

$$z_{k+1} = -\frac{1}{z_k}, \qquad z_{k+2} = -\frac{1}{z_k} + n,$$

where n is so determined that $-\tfrac{1}{2} \le x_{k+2} < \tfrac{1}{2}$. This gives

$$z_{k+2} = \frac{n x_k - 1 + i n y_k}{x_k + i y_k},$$

so that

$$|z_{k+2}|^2 = \frac{(n x_k - 1)^2 + n^2 y_k{}^2}{x_k{}^2 + y_k{}^2}.$$

If $|n| \ge 2$,

$$|z_{k+2}|^2 \ge \frac{4 \cdot \tfrac{1}{4}}{1} = 1,$$

while if $|n| = 1$, the hypothetical inequality $|z_{k+2}|^2 < 1$ gives

$$(x_k - n)^2 + y_k{}^2 < x_k{}^2 + y_k{}^2,$$

which says that z_k is farther from the origin than from the point $z = n$, which is false from the first inequality of (11). Finally, if $n = 0$, then $|z_{k+2}|^2 = 1/|z_k|^2 > 1$. Hence in all cases, $|z_{k+2}| \ge 1$, and $-\tfrac{1}{2} \le \operatorname{Re} z_{k+2} < \tfrac{1}{2}$. If z_{k+2} is still not in R (which may happen if $|z_{k+2}| = 1$) then $W z_{k+2}$ is in R, and the proof is complete.

Moreover, the proof has shown that S and W are generators of Γ, since every transformation of Γ can be written in the form

$$S_k{}^{n_k} W \cdots W S_2{}^{n_2} W S_1{}^{n_1}.$$
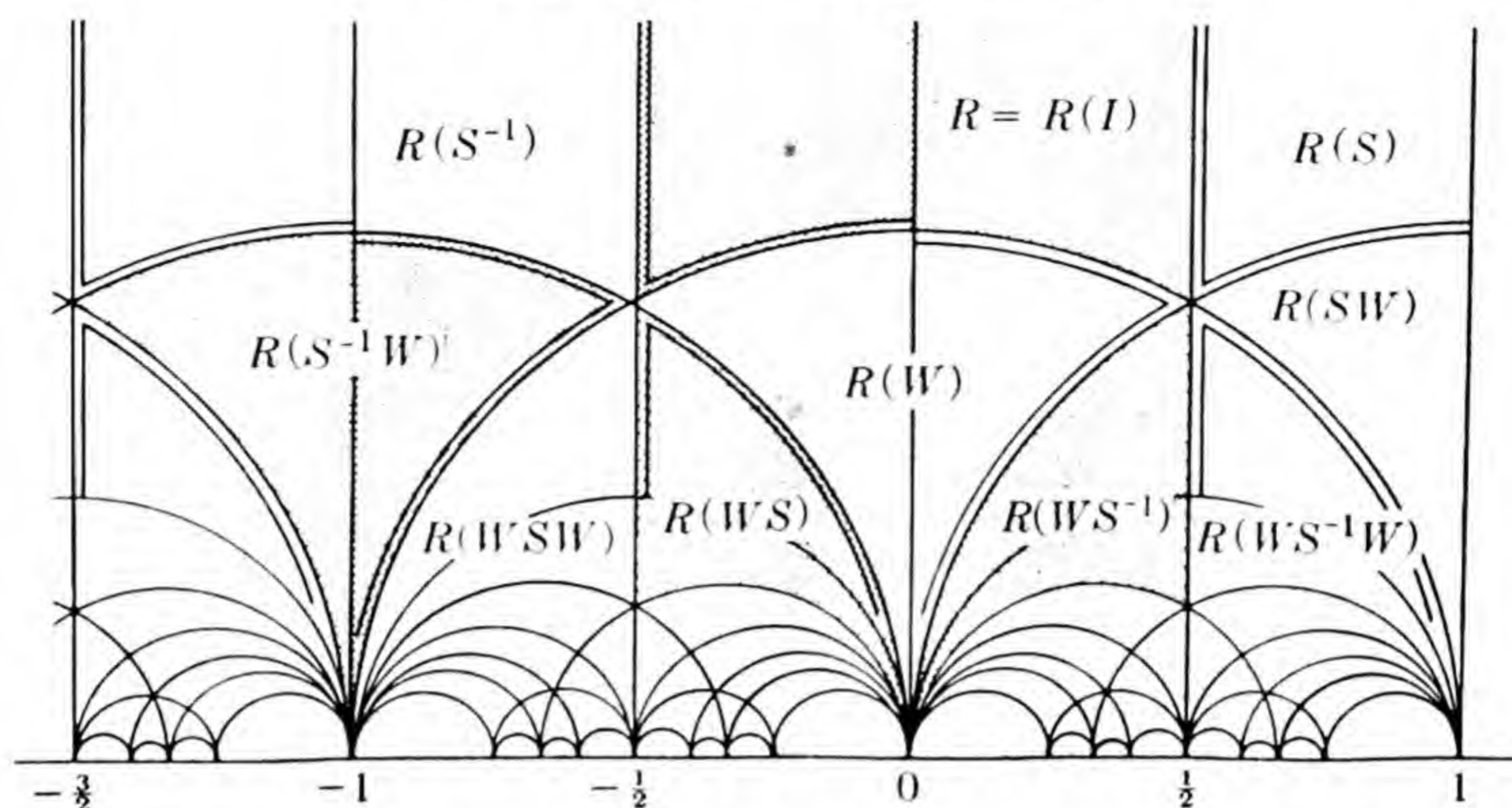
FIGURE 1–2

A geometric representation of the group Γ is given in Fig. 1–2. Here we have considered the region R as the image of itself under the identity transformation I, and have put $R = R(I)$. The congruent unshaded region to the left of R is then $R(S^{-1})$, in the sense that if a point z' of it is equivalent to a point z of R, then $z' = S^{-1}z$. To put it differently, S^{-1} maps R onto this region, just as W maps R onto the unshaded region $R(W)$ immediately below R. The semicircular arcs are portions of the circles $C(T)$; infinitely many of them terminate in each rational point on the real axis. If the drawing and shading were completed, any shaded or unshaded region could be taken as a fundamental region. Each fundamental region or "double triangle" is bounded by three arcs, with vertex angles of 0, $\pi/3$, and $\pi/3$. The heavy arc inside each region indicates the portion of the boundary which is to be included in the region.

PROBLEMS

1. Find the point in R to which the point

$$-\frac{3 + 2i}{8 + 6i}$$

is equivalent, by the method used in the proof of Theorem 1–1. Do you see an easier way, for this particular number?

2. If the term "circle" is used in the broad sense to include straight lines, show that the transformations of Γ send circles into circles. Under what circumstances are the image circles actually lines? What can be said about such a line if, for every point z on the original circle, Im $z \geq 0$?

3. Verify that, in the notation of the text, $(SW)^3 = I$.

4. Show that the transformations

$$P: \quad z' = 1 - z, \qquad Q: \quad z' = \frac{1}{z}$$

generate a group of six elements (the group of anharmonic ratios) of which a fundamental region is the set of points z such that

$$\operatorname{Im} z > 0, \qquad |z| > 1, \qquad \text{and} \qquad |z - 1| > 1,$$

together with half the boundary of this region, leading from $(1 + i\sqrt{3})/2$ to infinity in one direction. Sketch the analog of Fig. 1-2 for this group. [Note that the transformations of the group do not carry U into itself.]

1-4 Reduced definite forms. With the help of the facts now known about the modular group, we can deal with the question of how to decide whether two given binary quadratic forms are equivalent. We must consider separately the essentially different cases in which the discriminant $\Delta = 4ac - b^2$ of the form $ax^2 + bxy + cy^2$ is positive or negative. (We put aside the degenerate case in which $\Delta = 0$.) If $\Delta > 0$, the form is called *definite*, otherwise *indefinite*. The definite forms can be further classified as *positive* or *negative*, according as $a > 0$ or $a < 0$. The reason for this terminology is that the polynomial $az^2 + bz + c$ associated with a definite form has nonreal zeros, so that the form

$$y^2 \left\{ a \left(\frac{x}{y} \right)^2 + b \frac{x}{y} + c \right\}$$

has the same sign as a for every choice of x and y except $x = y = 0$, while an indefinite form can have values of either sign. We shall first consider definite forms, restricting our attention to positive forms, since the treatment of negative forms is almost identical.

Since the matrix of a form is a little cumbersome, we shall use the symbol $[a, b, c]$ to designate the form $ax^2 + bxy + cy^2$. It is to be clearly understood that this is simply an abbreviation, and cannot be combined with like symbols as matrices can.

Let us consider then a positive definite form $f(x, y) = [a, b, c]$, in which $\Delta > 0$, $a > 0$, and $c > 0$. For the time being, we do not require that a, b, and c be integers. Then the quadratic polynomial

$f(z) = az^2 + bz + c$ has zeros

$$\frac{-b \pm \sqrt{-\Delta}}{2a} \; ;$$

of these, we single out the one with positive imaginary part and call it ω. Thus to the form $[a, b, c]$ there corresponds the point ω in the upper half plane. Conversely, each point in U corresponds to exactly one form of discriminant Δ. For if z_0 is such a point, and $\bar{z}_0$ is its complex conjugate, then there is a unique number $\varkappa$ such that the quadratic expression $\varkappa(z - z_0)(z - \bar{z}_0)$ has discriminant Δ. Hence if we consider only forms of given discriminant Δ (which is all that is required in the equivalence problem, since equivalent forms have the same discriminant), there is a one-to-one correspondence between points of U and forms of that discriminant. Moreover, if the points ω_1 and ω_2 are associated with the forms f_1 and f_2 of discriminant Δ, and if a transformation T of Γ carries f_1 into f_2, then it carries ω_1 into ω_2. It therefore makes no difference whether one speaks of the form f or the point ω, as far as the operations of Γ are concerned. We call ω the *representative* of f.

It should now be clear how to decide whether or not two forms are equivalent. If they do not have the same discriminant, they are not equivalent. If they have, they are equivalent if and only if their representatives are equivalent, and this can be decided by transforming the representatives into the fundamental region R, where they must be identical to be equivalent. This leads us to define a *reduced form* as one whose representative is in R; *reduced forms are equivalent if and only if they are identical, and each class of equivalent forms contains exactly one reduced form.*

Since

$$\omega = \frac{-b + \sqrt{-\Delta}}{2a} = \frac{-b}{2a} + i \frac{\sqrt{\Delta}}{2a},$$

$$|\omega|^2 = \frac{b^2 + \Delta}{4a^2} = \frac{c}{a},$$

(12)

ω is in R if and only if $-\frac{1}{2} \leq -b/2a < \frac{1}{2}$, and either $c/a > 1$, or $c/a = 1$ and $-\frac{1}{2} \leq -b/a \leq 0$. Simplifying, we have that $[a, b, c]$ is reduced if and only if either

$$-a < b \leq a < c \qquad \text{or} \qquad 0 \leq b \leq a = c. \qquad (13)$$

Prove the assertion, made in the text, that if ω_1 and ω_2 are the representatives of the forms f_1 and f_2 with discriminant Δ, and if a T in Γ carries f_1 into f_2, it carries ω_1 into ω_2.

1–5 Reduction of definite forms. A given form can be transformed into its equivalent reduced form by exactly the process used in the proof that R is a fundamental region of Γ. That is, by a translation S^{n_1}, ω can be changed into ω', where $-\frac{1}{2} \le \text{Re }\omega' < \frac{1}{2}$; if ω' is not in R, we begin afresh with $W\omega'$, etc. The translation $z' = z + n_1$ must be such that

$$-\frac{1}{2} \le -\frac{b}{2a} + n_1 < \frac{1}{2},$$

or

$$b = 2an_1 + r_1,$$

where $-a < r_1 \le a$. The transformation $z' = z + n_1$ has matrix

$$S^{n_1} = \begin{pmatrix} 1 & n_1 \\ 0 & 1 \end{pmatrix},$$

but we must now revert to the inverse transformation $z = z' - n_1$ to utilize the results of Section 1–1, which were based on the equations (2). If we put

$$M = \begin{pmatrix} 1 & -n_1 \\ 0 & 1 \end{pmatrix} = S^{-n_1},$$

then, as we saw earlier, M carries a form with matrix F into one with matrix

$$G = \overline{M}FM,$$

so that in this case, if we let the result of the first translation be $f_1(x, y) = XF_1\bar{X}$, then

$$F_1 = \begin{pmatrix} 1 & 0 \\ n_1 & 1 \end{pmatrix} F \begin{pmatrix} 1 & -n_1 \\ 0 & 1 \end{pmatrix}.$$

Similarly, if F_2 is the result of applying the inversion W to F_1, then

$$F_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} F_1 \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

A simple calculation shows that, if $f_1 = [a, b, c]$, then $f_2 = [c, -b, a]$.

Thus we have the following algorithm for reducing $f = [a, b, c]$: find n_1 and r_1 such that

$$b = 2an_1 - r_1, \qquad -a \leq r_1 < a,$$

and compute $f_1 = [a_1, b_1, c_1]$, where

$$F_1 = \begin{pmatrix} 1 & 0 \\ -n_1 & 1 \end{pmatrix} F \begin{pmatrix} 1 & -n_1 \\ 0 & 1 \end{pmatrix},$$

so that $f_1 = [a, b - 2an_1, n_1{}^2 a - bn_1 + c]$. If f_1 is not reduced, put $f_2 = [c_1, -b_1, a_1] = [a_2, b_2, c_2]$. If f_2 is not reduced, repeat the entire procedure. For some k, f_k will be reduced.

The discussion thus far has been valid for positive definite forms with arbitrary real coefficients. For the remainder of this section and the next, we consider only *integral* forms, that is, those with integral coefficients.

THEOREM 1-3. *There are only finitely many classes of integral definite forms of given discriminant.*

Proof: To each class there belongs just one reduced form $[a, b, c]$ satisfying the conditions (13). Since

$$4a^2 \leq 4ac = \Delta + b^2 \leq \Delta + a^2,$$

the inequality $0 < a \leq \sqrt{\Delta/3}$ holds for each reduced form, so that there are only finitely many possible values of a for fixed Δ. Since $|b| \leq a$, the same is true of b, and for each pair a, b there is at most one integer c such that $4ac - b^2 = \Delta$.

If, for example, $\Delta = 3$, then $0 < a \leq 1$, so that $a = 1$ and hence $b = 0$ or 1; from this it is easily seen that the only integral reduced form of discriminant 3 is $x^2 + xy + y^2$. There is also just one class of discriminant 4, and its reduced form is $x^2 + y^2$.

PROBLEMS

1. Find all reduced integral definite forms of discriminant $\Delta \leq 20$.
2. Find the reduced form equivalent to $[117, 103, 100]$.

1–6 Representations by definite forms. If a transformation of Γ leaves a quadratic form unchanged, it is called an *automorph* of the form. Since an automorph also leaves the representative of the form unchanged, and is the only kind of transformation which does, the following theorem is an easy consequence of Theorem 1–2.

THEOREM 1–4.   *The only automorphs of $a(x^2 + y^2)$ are*

$$\begin{cases} x = \pm x', \\ y = \pm y', \end{cases} \quad and \quad \begin{cases} x = \pm y', \\ y = \mp x'. \end{cases}$$

The only automorphs of $a(x^2 + xy + y^2)$ are

$$\begin{cases} x = \pm x', \\ y = \pm y', \end{cases} \quad \begin{cases} x = \mp y', \\ y = \pm x' \pm y', \end{cases} \quad and \quad \begin{cases} x = \pm x' \pm y', \\ y = \mp x'. \end{cases}$$

Any positive reduced form distinct from these two has only the automorphs

$$\begin{cases} x = \pm x', \\ y = \pm y'. \end{cases}$$

An integer n is said to be *properly representable* by an integral form $[a, b, c]$ of discriminant Δ if there are relatively prime integers α, γ such that $a\alpha^2 + b\alpha\gamma + c\gamma^2 = n$. For such α, γ, there are β_0 and δ_0 such that $\alpha\delta_0 - \beta_0\gamma = 1$, and, in fact,

$$\alpha\delta - \beta\gamma = 1,$$

if, for some integer t,

$$\beta = \beta_0 + \alpha t,$$

$$\delta = \delta_0 + \gamma t.$$

If we make the substitution

$$x = \alpha x' + \beta y',$$
$$y = \gamma x' + \delta y', \tag{14}$$

then $[a, b, c]$ goes into a form $[n, m, l]$ with first coefficient n, by equations (5). Also by (5),

$$m = 2a\alpha(\beta_0 + \alpha t) + b(\alpha\delta_0 + \alpha\gamma t + \beta_0\gamma + \alpha\gamma t) + 2c\gamma(\delta_0 + \gamma t)$$
$$= 2a\alpha\beta_0 + b(\alpha\delta_0 + \beta_0\gamma) + 2c\gamma\delta_0 + 2nt,$$

so that m is determined modulo $2n$. Choose m so that $0 \leq m < 2n$; then t is fixed, β and δ are unique, and l is determined by the discriminant:

$$4ln - m^2 = \Delta.$$

THEOREM 1–5.   *Let α, γ be a proper representation of $n > 0$ by the integral form $[a, b, c]$ of discriminant Δ. Then there are unique integers β and δ such that $\alpha\delta - \beta\gamma = 1$, and the substitution* (14)

replaces $[a, b, c]$ *by the equivalent form* $[n, m, l]$, *where* $0 \leq m < 2n$, *m satisfies the congruence*

$$m^2 \equiv -\Delta \pmod{4n}, \tag{15}$$

and

$$l = \frac{m^2 + \Delta}{4n}. \tag{16}$$

Thus to each proper representation of n by $[a, b, c]$ there corresponds a unique form which has first coefficient n and satisfies certain auxiliary conditions. The appropriate converse, which we now consider, gives the number of such representations, and provides an effective method of finding them. If m is a solution of (15) and $0 \leq m < 2n$, then $4n - m$ is also a root, and $2n \leq 4n - m < 4n$. We shall refer to m as a *minimum root* if $0 \leq m < 2n$.

THEOREM 1–6. *Let* $w(f)$ *be the number of automorphs of* $f = [a, b, c]$, *an integral positive form of discriminant* Δ. *Let* n *be a positive integer. Corresponding to each minimum root* m *of the congruence* (15), *determine* l *by equation* (16). *Then the number of proper representations of* n *by* f *is* $w(f)$ *times the number of such forms* $[n, m, l]$ *which are equivalent to* f. *In particular, if there is only one class of discriminant* Δ, *the number of proper representations is* $w(f)$ *times the number of minimum roots of* (15).

Proof: Suppose that $g = [n, m, l]$ is a form of the type described in the theorem. Then if f is not equivalent to g, Theorem 1–5 shows that there is no representation of n by f corresponding to the minimum root m. If f is equivalent to g, let T be the matrix of a substitution which replaces f by g, and let A be the matrix of an automorph of f. Then

$$G = \bar{T}FT \quad \text{and} \quad F = \bar{A}FA,$$

so

$$(\overline{AT})F(AT) = \bar{T}\bar{A}FAT = \bar{T}FT = G,$$

so that AT is also the matrix of a substitution which carries f into g. Conversely, if for any U,

$$G = \bar{U}FU,$$

then $\bar{U}FU = \bar{T}FT$, and

$$F = \bar{T}^{-1}\bar{U}FUT^{-1},$$

so that UT^{-1} is the matrix A of an automorph of f, and $U = AT$. Hence there are exactly $w(f)$ substitutions which replace f by g.

If
$$T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix},$$

and f has only two automorphs (see Theorem 1–4), then

$$AT = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -\alpha & -\beta \\ -\gamma & -\delta \end{pmatrix},$$

and α, γ and $-\alpha, -\gamma$ give two distinct proper representations, since $(\alpha, \gamma) = 1$ and therefore α and γ are not both zero. If $f \sim a(x^2 + y^2)$, then

$$AT = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -\alpha & -\beta \\ -\gamma & -\delta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} -\gamma & -\delta \\ \alpha & \beta \end{pmatrix}$$

$$\text{or} \quad \begin{pmatrix} \gamma & \delta \\ -\alpha & -\beta \end{pmatrix},$$

and the representations α, γ; $-\alpha, -\gamma$; $-\gamma, \alpha$; and $\gamma, -\alpha$ are again distinct. If $f \sim a(x^2 + xy + y^2)$, then AT is one of the matrices

$$\pm \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \quad \pm \begin{pmatrix} -\gamma & -\delta \\ \alpha + \gamma & \beta + \delta \end{pmatrix} \quad \text{or} \quad \pm \begin{pmatrix} \alpha + \gamma & \beta + \delta \\ -\alpha & -\beta \end{pmatrix},$$

and these also lead to distinct representations.

If there is only one class of discriminant Δ, then f and g are necessarily equivalent, so that all minimum roots of (15) lead to representations. The proof is complete.

In the case of *primitive* forms (those having relatively prime coefficients), $w(f)$ depends only on Δ: $w(f) = 6, 4,$ or 2 according as Δ is $3, 4,$ or larger than 4. If $f(x, y) = x^2 + y^2$, so that $\Delta = 4$, then m must be even to satisfy (15). Let $m = 2m_1$; then $m_1^2 \equiv -1 \pmod{n}$, and $0 \le m < 2n$ means $0 \le m_1 < n$, so that the number of proper representations of n as a sum of two squares is four times the number of solutions of the congruence $u^2 \equiv -1 \pmod{n}$. This result was obtained in Theorem 7–5, Volume I, by quite different methods.

PROBLEMS

1. Find β, δ, m, l of Theorem 1–5 corresponding to the proper representation 3, 5 of 118 by $[2, -5, 7]$.

2. What is the number of proper representations of 28 by $[1, 1, 2]$? Find them.

3. Use Theorem 1–6 to discuss the proper representability of 10 by $[2, 1, 2]$.

4. Show that every prime congruent to 1 or 3 (mod 8) has a unique proper representation in the form $x^2 + 2y^2$ with $x > 0$, $y > 0$. More generally, show that if n is the product of powers of r such primes, then n has 2^{r+1} proper representations in this form.

1–7 Indefinite forms. The behavior of indefinite binary forms is remarkably different from that of forms with positive discriminant. For example, any integral indefinite form whose discriminant is not the negative of a square has infinitely many automorphs, and therefore represents any integer in infinitely many ways if it represents it at all. Moreover, there seems to be no natural way to pick out a unique reduced form in each equivalence class, although we shall find a finite set of canonical forms in the case of integral forms.

Hereafter we restrict attention to integral forms $[a, b, c]$, and put

$$D = -\Delta = b^2 - 4ac > 0.$$

If D is a square, then $[a, b, c]$ factors into two linear factors with integral coefficients. We dismiss this degenerate case, and hereafter require that D be a nonsquare integer. Finally, for the sake of simplicity we consider only the case that $[a, b, c]$ is primitive. We see from equations (5) (proof by contradiction) that any form $[A, B, C]$ equivalent to a primitive form is again primitive.

As before, there is associated with $[a, b, c]$ the quadratic equation

$$az^2 + bz + c = 0,$$

which this time has two real roots, say

$$\omega_1 = \frac{-b + \sqrt{D}}{2a}, \qquad \omega_2 = \frac{-b - \sqrt{D}}{2a}.$$

It is easily verified that a transformation of the modular group which sends $[a, b, c]$ into $[a', b', c']$ sends ω_1 into ω_1' and ω_2 into ω_2', and never ω_1 into ω_2'. We call ω_1 the *first root*, and ω_2 the *second root*.

As C. Hermite noticed, there is also associated with

$$[a, b, c] = a(x - \omega_1 y)(x - \omega_2 y)$$

a family of definite forms

$$\varphi_t(x, y) = \frac{a}{2t}(x - \omega_1 y)^2 + \frac{at}{2}(x - \omega_2 y)^2,$$

where $t > 0$ is a real parameter. A simple calculation shows that the discriminant of $\varphi_t(x, y)$ is D, for every $t > 0$. Reverting to the quotient variable $z = x/y$, we find the zeros of $\varphi_t(z)$ to be those points z_t such that

$$\frac{1}{t}(z_t - \omega_1)^2 = -t(z_t - \omega_2)^2,$$

or

$$z_t - \omega_1 = \pm it(z_t - \omega_2).$$

The transformation $z' = iz$ rotates the plane about the origin through the angle $\pi/2$; it follows from the last equation that the line segment connecting z_t with ω_1 is perpendicular to the segment connecting z_t with ω_2, and hence that z_t lies on the circle having as diameter the segment which connects ω_1 and ω_2. If, as usual, we take that root z_t which has positive imaginary part as the representative of φ_t, then we have associated with $[a, b, c]$ the semicircle Σ in U connecting ω_1 and ω_2. As t varies from 0^+ to ∞, z_t describes Σ from ω_1 to ω_2; we can think of the semicircle as oriented with this sense, inasmuch as the orientation is preserved under transformations of Γ. This orientation is necessary, since otherwise there would be no way of distinguishing the (usually inequivalent) forms $[a, b, c]$ and $[-a, -b, -c]$. The form is now completely described by specifying its oriented semicircle Σ and its discriminant $-D$.

An indefinite form f will be called *reduced* if the associated semi-circle intersects the fundamental region R considered earlier. Thus f is reduced if and only if the definite form φ_t is reduced for some t. The fact that any indefinite form is equivalent to a reduced form is an immediate consequence of the fact that φ_1, for example, is equivalent to a reduced definite form: the transformation which carries φ_1 into a reduced form also carries f into a reduced form. The difficulty lies in showing that each indefinite integral form is equivalent to only finitely many reduced forms. To do this, we must first examine an important subgroup of Γ which is intimately connected with f.

1-8 The automorphs of indefinite forms. A transformation of Γ which leaves $[a, b, c]$ unchanged also leaves ω_1 and ω_2 fixed. The fixed points of the transformation

$$z' = \frac{\alpha z + \beta}{\gamma z + \delta}$$

are those points ω such that

$$\omega = \frac{\alpha\omega + \beta}{\gamma\omega + \delta},$$

or

$$\gamma\omega^2 + (\delta - \alpha)\omega - \beta = 0. \tag{17}$$

Suppose that the roots of this equation are ω_1 and ω_2. These numbers are also the roots of the equation $a\omega^2 + b\omega + c = 0$; since $(a, b, c) = 1$, it follows that for some integer u,

$$\gamma = au, \tag{18}$$

$$\delta - \alpha = bu,$$

$$-\beta = cu. \tag{19}$$

Putting $\delta + \alpha = t$, we have

$$\alpha = \frac{t - bu}{2}, \qquad \delta = \frac{t + bu}{2}, \tag{20}$$

where t and u are such that

$$1 = \alpha\delta - \beta\gamma = \frac{t^2 - b^2 u^2}{4} + acu^2 = \frac{t^2 - Du^2}{4},$$

or

$$t^2 - Du^2 = 4. \tag{21}$$

Conversely, if t and u are solutions of (21), and α, β, γ, and δ are determined by equations (18) through (20), then (17) reduces to $u(a\omega^2 + b\omega + c) = 0$ and $\alpha\delta - \beta\gamma = 1$. This proves

THEOREM 1-7.   *The set of all automorphs of the primitive indefinite form $[a, b, c]$ is given by the set of all matrices*

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

with α, β, γ, and δ determined by equations (18), (19), and (20), where t and u run over the integral solutions of the Pell equation (21).

Originally, automorphs were defined as substitutions giving z in terms of z', while we have here used the inverse transformation giving z' in terms of z. But if $F = \bar{A}FA$, then $F = \bar{A}^{-1}FA^{-1}$, so that the inverse of an automorph is also an automorph, and the set of all automorphs coincides with the set of all inverse automorphs. This fact has much greater significance than in its application above. For since the product of two automorphs is again an automorph, the automorphs of f form a subgroup of Γ, which we shall designate by $\Gamma_A(f)$. (The elements of $\Gamma_A(f)$ will be taken sometimes as transformations and sometimes as their matrices. The ambiguity resulting from the fact that the matrices A and $-A$ correspond to different substitutions in the form but to the same fractional transformation of Γ should cause no difficulty if the reader remains aware of it.) Using well-known properties of the solutions of Pell's equation,* $\Gamma_A(f)$ can be characterized as follows.

THEOREM 1–8. *$\Gamma_A(f)$ is the infinite cyclic group generated by the matrix*

$$V = \begin{pmatrix} \frac{1}{2}(t_0 - bu_0) & -cu_0 \\ au_0 & \frac{1}{2}(t_0 + bu_0) \end{pmatrix};$$

if A is any automorph of f, then $A = V^n$ for some integer n, positive, negative or zero. Here t_0, u_0 is the minimal positive solution of equation (21).

(The ambiguity mentioned above is exemplified here: every transformation $z' = (\alpha z + \beta)/(\gamma z + \delta)$ of $\Gamma_A(f)$ can be made to have matrix V^n, but the set of all substitutions which leave f fixed is given by $X = \pm X'V^n$.)

Proof: According to Theorem 1–7, $\Gamma_A(f)$ is the group of matrices

$$\begin{pmatrix} \frac{1}{2}(t - bu) & -cu \\ au & \frac{1}{2}(t + bu) \end{pmatrix}, \qquad t^2 - Du^2 = 4,$$

so it is to be shown that each of these matrices is a power of V.
If we put

$$\tfrac{1}{2}(t_0 + u_0\sqrt{D})^n = \tfrac{1}{2}(t_n + u_n\sqrt{D})$$

———

*Pell's equation is discussed in Volume I, Chapter 8. The minimal positive solution is described in Theorem 8–7.

for each n, then

$$\tfrac{1}{2}(t_{n+1} + u_{n+1}\sqrt{D}) = \tfrac{1}{4}(t_n + u_n\sqrt{D})(t_0 + u_0\sqrt{D})$$
$$= \tfrac{1}{4}(t_0 t_n + D u_0 u_n) + \tfrac{1}{4}(t_0 u_n + t_n u_0)\sqrt{D},$$

so that

$$t_{n+1} = \tfrac{1}{2}(t_0 t_n + D u_0 u_n), \qquad u_{n+1} = \tfrac{1}{2}(t_n u_0 + t_0 u_n).$$

Now suppose that

$$V^{n+1} = \begin{pmatrix} \tfrac{1}{2}(t_n - b u_n) & -c u_n \\ a u_n & \tfrac{1}{2}(t_n + b u_n) \end{pmatrix},$$

an assumption which is correct for $n = 0$. Then

$$V^{n+2} = V V^{n+1} = \begin{pmatrix} \tfrac{1}{2}(t_0 - b u_0) & -c u_0 \\ a u_0 & \tfrac{1}{2}(t_0 + b u_0) \end{pmatrix}$$
$$\times \begin{pmatrix} \tfrac{1}{2}(t_n - b u_n) & -c u_n \\ a u_n & \tfrac{1}{2}(t_n + b u_n) \end{pmatrix}$$
$$= \begin{pmatrix} \tfrac{1}{2}(t_{n+1} - b u_{n+1}) & -\tfrac{1}{2}c(u_0 t_n + u_n t_0) \\ \tfrac{1}{2}a(u_0 t_n + u_n t_0) & \tfrac{1}{2}(t_{n+1} + b u_{n+1}) \end{pmatrix}$$
$$= \begin{pmatrix} \tfrac{1}{2}(t_{n+1} - b u_{n+1}) & -c u_{n+1} \\ a u_{n+1} & \tfrac{1}{2}(t_{n+1} + b u_{n+1}) \end{pmatrix},$$

and by induction, V^n is of the supposed form for all $n \geq 0$. Similarly, it can be shown that

$$V^n = \begin{pmatrix} \tfrac{1}{2}(t_{n-1} - b u_{n-1}) & -c u_{n-1} \\ a u_{n-1} & \tfrac{1}{2}(t_{n-1} + b u_{n-1}) \end{pmatrix},$$

so that V^n is also of the supposed form for all $n \leq 0$. Hence the matrix corresponding to any solution of equation (21) is a power of V, and the theorem is proved.

As usual, it is useful to know a fundamental region of $\Gamma_A(f)$.

THEOREM 1–9. *Suppose that the perpendicular bisector C_0 of the segment l joining ω_1 and ω_2 is mapped by V into the circle C_1. Then C_1 does not intersect C_0, and the (infinite) region between them, together with C_1, ω_1, and ω_2, is a fundamental region of $\Gamma_A(f)$.*

Proof: If the arbitrary transformation

$$z' = T(z) = (\alpha z + \beta)/(\gamma z + \delta)$$

has the distinct fixed points z_1 and z_2, then by dividing $z' - z_1$ by $z' - z_2$ we get

$$\frac{z' - z_1}{z' - z_2} = \frac{\alpha z + \beta - z_1(\gamma z + \delta)}{\alpha z + \beta - z_2(\gamma z + \delta)} = \frac{(\alpha - \gamma z_1)z + (\beta - \delta z_1)}{(\alpha - \gamma z_2)z + (\beta - \delta z_2)}$$

$$= \frac{\alpha - \gamma z_1}{\alpha - \gamma z_2} \cdot \frac{z + (\beta - \delta z_1)/(\alpha - \gamma z_1)}{z + (\beta - \delta z_2)/(\alpha - \gamma z_2)}$$

$$= \frac{\alpha - \gamma z_1}{\alpha - \gamma z_2} \cdot \frac{z - T^{-1}(z_1)}{z - T^{-1}(z_2)} = \frac{\alpha - \gamma z_1}{\alpha - \gamma z_2} \cdot \frac{z - z_1}{z - z_2}.$$

In the case at hand, T is the transformation

$$V: \qquad z' = \frac{\frac{1}{2}(t_0 - bu_0)z - cu_0}{au_0 z + \frac{1}{2}(t_0 + bu_0)}$$

with fixed points ω_1 and ω_2, and

$$\frac{\alpha - \gamma z_1}{\alpha - \gamma z_2} = \frac{\frac{1}{2}(t_0 - bu_0) - au_0(-b + \sqrt{D})/2a}{\frac{1}{2}(t_0 - bu_0) - au_0(-b - \sqrt{D})/2a} = \frac{t_0 - \sqrt{D}\, u_0}{t_0 + \sqrt{D}\, u_0}.$$

We put

$$K = \frac{t_0 - \sqrt{D}\, u_0}{t_0 + \sqrt{D}\, u_0},$$

and have for V the representation

$$\frac{z' - \omega_1}{z' - \omega_2} = K \frac{z - \omega_1}{z - \omega_2}. \tag{22}$$

It follows that V^n is the same transformation with K replaced by K^n; this could be used to give a second proof of Theorem 1–8.

By its definition, K is a real number between 0 and 1. Since the perpendicular bisector C_0 of the segment l joining ω_1 and ω_2 has the equation $|z - \omega_1| = |z - \omega_2|$, V^n transforms it into $|z' - \omega_1| = K^n|z' - \omega_2|$, as we see by taking absolute values in (22). If we put $z = x + iy$, the last equation becomes

$$C_n: \qquad (x - \omega_1)^2 + y^2 = K^{2n}((x - \omega_2)^2 + y^2), \qquad n \gtreqless 0$$

and it is a matter of simple analytic geometry to prove the following assertions: for positive n, C_n is a circle with its center on the real axis, on the extension through ω_1 of l; it contains ω_1 in its interior;

it lies entirely on that side of C_0 on which ω_1 lies; and its radius approaches zero as n increases. For negative n, the circles C_n lie on the other side of C_0, contain ω_2, and close down on ω_2 as $|n|$ increases. Some of these circles are shown in Fig. 1–3. The lightly shaded region $R_A(1)$, which is the region described in Theorem 1–9, is the set of points z such that

$$K \leq \left| \frac{z - \omega_1}{z - \omega_2} \right| < 1,$$

and it is clearly transformed by V into the set $R_A(V)$ of points z such that

$$K^2 \leq \left| \frac{z - \omega_1}{z - \omega_2} \right| < K,$$
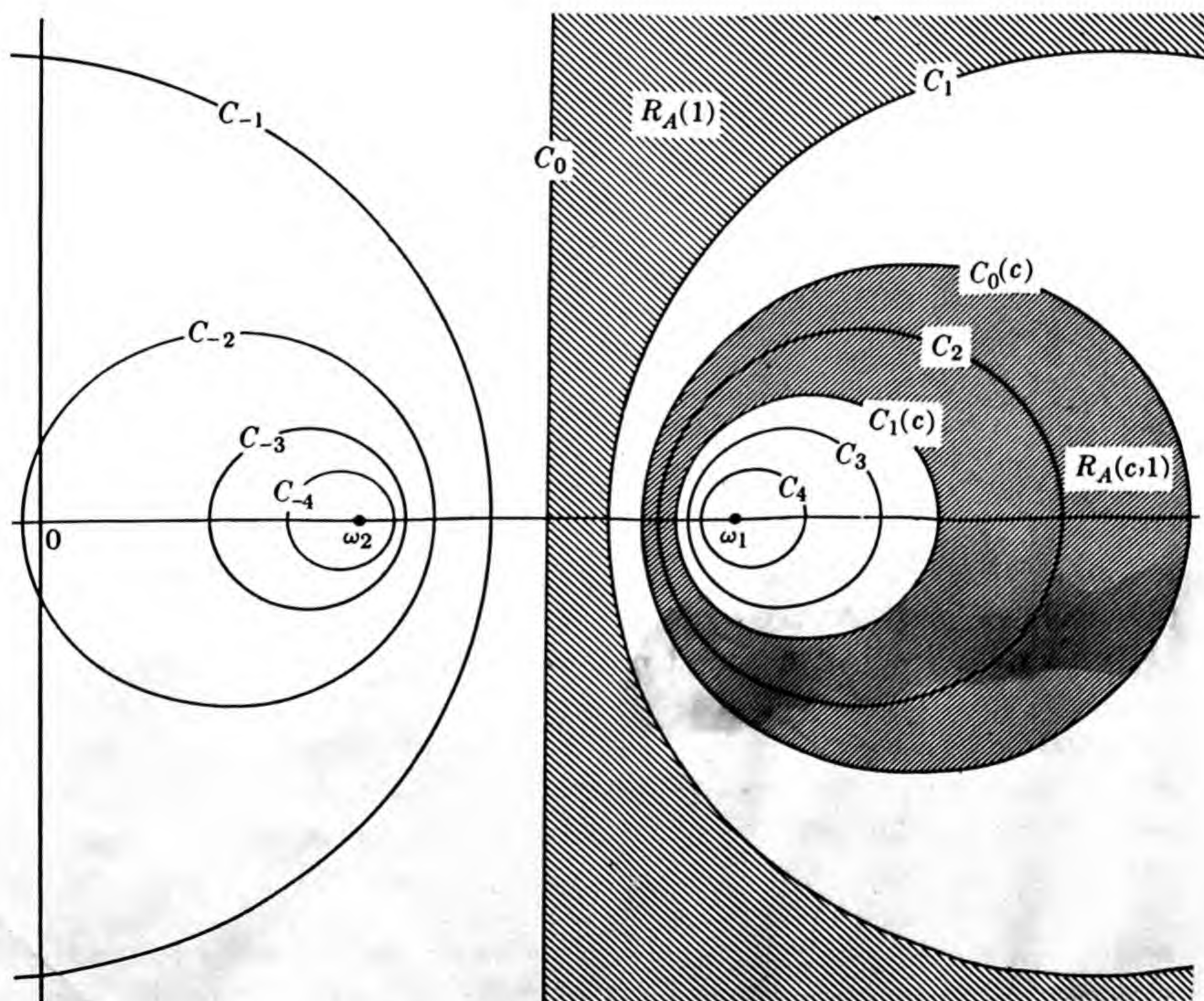
FIGURE 1–3

which is the region between C_1 and C_2, including C_2. In general, V^n transforms $R_A(1)$ into the region $R_A(V^n)$ between C_n and C_{n+1}, including C_{n+1}. Since the entire plane, excluding ω_1 and ω_2, is covered in this fashion, and no point is in two such regions, any one of them, together with ω_1 and ω_2, is a fundamental region of $\Gamma_A(f)$, and the proof is complete.

We are concerned here only with the upper half-plane U; relative to this, a fundamental region of $\Gamma_A(f)$ is that portion of any one of the above regions which lies in U.

In the next section it will be convenient to have slightly more freedom in choosing a fundamental region of $\Gamma_A(f)$. We get this by noticing that, instead of beginning with the line C_0, we could have started with any member of the family of circles

$$\left| \frac{z - \omega_1}{z - \omega_2} \right| = c. \tag{23}$$

For fixed $c > 0$, a fundamental region $R_A(c, 1)$ would then be the ring between the circle (23), which we might call $C_0(c)$, and its transform

$$C_1(c): \qquad \left| \frac{z - \omega_1}{z - \omega_2} \right| = Kc;$$

the argument given above carries through with no change except for the introduction of a factor c in certain equations. Such a region is shown heavily shaded in Fig. 1–3.

1–9 Reduction of indefinite forms. The semicircle Σ representing the form f is the upper half of the circle given parametrically by

$$\left(\frac{z - \omega_1}{z - \omega_2} \right)^2 = -t^2, \qquad 0 \le t \le \infty.$$

The generating automorph V, given by equation (22), changes Σ into the upper half of the circle

$$\left(\frac{z' - \omega_1}{z' - \omega_2} \right)^2 = -Kt^2 = -(t\sqrt{K})^2, \qquad 0 \le t \le \infty,$$

which is the same circle with a different parameter. In other words, Σ is transformed into itself by V, and hence by any element of $\Gamma_A(f)$,

in the sense that each point of Σ goes into some other point of Σ, although no points of Σ remain fixed except ω_1 and ω_2. In fact, that arc of Σ which lies in a fundamental region $R_A(c, 1)$ is mapped by V^n onto the arc of Σ which lies in the region $R_A(c, V^n)$, so that these various arcs are equivalent with respect to $\Gamma_A(f)$. Hence they are also equivalent with respect to the larger group Γ.

Now imagine Σ drawn in Fig. 1–3. For suitable choice of c, the circle $C_0(c)$ defined in the last section intersects Σ at a point on the boundary of one of the transforms of R, and this is then also true of the equivalent point which is the intersection of $C_1(c)$ and Σ. The arc between these two points is thus broken up by the boundaries of the double triangles in Fig. 1–2 into a finite number, say μ, of smaller arcs. If these short arcs are transformed back into R by suitable operations of Γ, then every point of Σ is equivalent to some point on each of these new arcs; in other words, there are precisely μ elements of Γ which transform Σ into a semicircle intersecting R. Hence there are precisely μ reduced forms equivalent to f.

THEOREM 1–10. *There are only finitely many reduced forms in any equivalence class of integral primitive indefinite forms.*

Using the definition of reduced form, it is simple to characterize reduced forms in terms of their coefficients. For clearly $[a, b, c]$ is reduced if and only if one or both of the points ρ and $-\rho^2$ are inside the semicircular region bounded by Σ, or if ρ is on Σ. The points below Σ in U are the points $z = x + iy$ such that

$$a(a(x^2 + y^2) + bx + c) < 0.$$

Since ρ and $-\rho^2$ have the coordinates

$$x = \pm \frac{1}{2}, \qquad y = \frac{\sqrt{3}}{2},$$

we have that f is reduced if and only if either

$$a(2a \pm b + 2c) < 0 \qquad \text{or} \qquad 2a - b + 2c = 0. \tag{24}$$

To find the set of reduced forms of the class containing a given form $[a, b, c]$, the procedure outlined for definite forms may first be used to reduce $\varphi_1(x, y) = a/2(x - \omega_1 y)^2 + a/2(x - \omega_2 y)^2 = ax^2 + bxy + (b^2 + D)y^2/4a$; the transformation which reduces $\varphi_1(x, y)$ also reduces $[a, b, c]$, say to $[a_1, b_1, c_1]$. Thus the semicircle Σ_1 represent-

ing $[a_1, b_1, c_1]$ intersects the fundamental region R of Γ, either in an arc or in the single point ρ. Starting from a point on Σ_1 in R, move along Σ_1 in the direction in which it is oriented. At the point at which Σ_1 leaves R, it enters one of the regions

$$R(S^{-1}),\ R(S^{-1}W),\ R(WSW),\ R(WS),\ R(W),$$

$$R(WS^{-1}),\ R(WS^{-1}W),\ R(SW),\ \text{or}\ R(S),$$

since these are the only regions adjacent to R (cf. Fig. 1–2). If it enters $R(T_1)$, then T_1^{-1} sends Σ_1 into a new semicircle Σ_2 (associated with $[a_2, b_2, c_2]$) which has an arc in R, and this arc is the image under T_1^{-1} of the portion of Σ_1 in $R(T_1)$. The same argument can now be applied to Σ_2, leading to a Σ_3 (associated with $[a_3, b_3, c_3]$) which has an arc in R, and this arc is the image under $T_2^{-1}T_1^{-1}$ of the arc of Σ_1 next encountered in moving along Σ_1 in the positive direction. If the process is repeated μ times, Σ_1 and $[a_1, b_1, c_1]$ will recur.

It is rather the exceptional case that Σ passes through ρ or $-\rho^2$. If it does not, the array of possible transformations listed above simplifies: the only T's to consider are then S, S^{-1}, and W. For example, consider the reduced form $[2, -4, -1]$, where $\omega_1 = (2 + \sqrt{6})/2$, $\omega_2 = (2 - \sqrt{6})/2$. Σ_1 goes from R to $R(W)$, so we make the inversion $W^{-1} = W$, or $z = -1/z'$. This replaces $[a, b, c]$ by $[c, -b, a]$, so here $[a_2, b_2, c_2] = [-1, 4, 2]$. Σ_2 goes from R to $R(S)$, so we make the translation S^{-1}, or $z = z' + 1$. In general this replaces $[a, b, c]$ by $[a, 2a + b, a + b + c]$, so here $[a_3, b_3, c_3] = [-1, 2, 5]$. Σ_3 also goes from R to $R(S)$, and we get $[a_4, b_4, c_4] = [-1, 0, 6]$. Σ_4 also goes from R to $R(S)$, and $[a_5, b_5, c_5] = [-1, -2, 5]$. A final application of S^{-1} gives $[a_6, b_6, c_6] = [-1, -4, 2]$. Since Σ_6 goes from R to $R(W)$, we invert, to get $[a_7, b_7, c_7] = [2, 4, -1]$. Σ_7 goes from R to $R(S^{-1})$, so we must make the translation $S: z = z' - 1$. In general, this replaces $[a, b, c]$ by $[a, b - 2a, a - b + c]$, so here $[a_8, b_8, c_8] = [2, 0, -3]$. A second application of S gives $[a_9, b_9, c_9] = [2, -4, -1] = [a_1, b_1, c_1]$, and we have the complete set of reduced forms for this class. If the algorithm were repeated indefinitely, a periodic sequence of forms would arise; it is therefore meaningful to speak of the *period* of reduced forms.

The following principle is useful in these calculations: If *after a translation* the inequality (24) is correct for just one choice of sign,

the next step is an inversion, while if it holds for both signs, the next step is a repetition of the translation. (S is never followed by S^{-1}, nor W by W.) The reason for this should become clear upon looking back at the derivation of (24).

THEOREM 1-11. *There are only finitely many classes of integral indefinite forms of given discriminant.*

Proof: First consider the primitive forms; for them it suffices to show that there are only finitely many reduced forms of given discriminant $\Delta = -D$. From (24) we get

$$2a^2 \pm ab \le -2ac,$$

so

$$4a^2 \pm 2ab + b^2 \le b^2 - 4ac = D.$$

But for each choice of sign, $4a^2 \pm 2ab + b^2$ is positive definite; it therefore represents only positive integers unless $a = b = 0$, and by Theorem 1-6, each of the integers $1, 2, \ldots, D$ is represented in only finitely many ways. Hence there are only finitely many choices for a and b, and for each choice, c is fixed by the requirement $b^2 = D + 4ac$. There are therefore only finitely many reduced forms, and hence only finitely many periods, and so only finitely many classes.

If a class contains an imprimitive form, say with $(a, b, c) = d$, then every form in that class also has divisor d, so that the class consists of the elements of a class of primitive forms with smaller D, each multiplied by d. There are only finitely many such classes.

PROBLEMS

1. Find the period of reduced forms belonging to the class of

$$x^2 + 7xy + 7y^2.$$

2. Show that Theorem 1-7 remains correct if the word "indefinite" is omitted, that is, if $D < 0$ (cf. Theorem 1-4).

3. Show that there is just one class of primitive forms with $D = 20$, and one class of imprimitive forms.

1-10 Representations. The discussion occurring be⋯ rems 1-4 and 1-6 made no use of the definiteness of the ⋯ therefore equally applicable to indefinite forms. Thus Theor⋯ ⋯ can be recast as follows.

THEOREM 1–12. *Let $f = [a, b, c]$ be a primitive integral indefinite form of discriminant Δ, where $D = -\Delta$ is not a square. Let n be an integer. Corresponding to each minimum root m of the congruence (15), determine l by (16). If none of the forms $[n, m, l]$ is equivalent to f, there are no proper representations of n by f. If at least one of the new forms is equivalent to f, there are infinitely many proper representations of n by f; they are given by the first columns of all the matrices AT, where A can be any automorph $\pm V^n$ of f, and T is any of a set of matrices which replace f by the various equivalent forms $[n, m, l]$, each form being obtained from just one T.*

PROBLEMS

1. Discuss the proper representation of 13 by $[1, 3, -1]$.
2. Show that the odd numbers properly represented by $x^2 + 4xy - y^2$ are those of the form

$$5^\epsilon \prod_{i=1}^{r} p_i^{\alpha_i},$$

where $\epsilon = 0$ or 1, $r \geq 0$, and $p_i \equiv \pm 1 \pmod{10}$ for $1 \leq i \leq r$ (cf. Problem 3, Section 1–10).

CHAPTER 2

ALGEBRAIC NUMBERS

2–1 Introduction. With a few exceptions, the theory developed up to this point, both in this volume and in the preceding introductory volume, has been self-contained, in the sense that the problems, which had to do with the ordinary integers, were solved without going outside this system. When considering the distribution of primes and the theory of quadratic forms, we made use of the real and complex numbers, but not in an intrinsically arithmetic fashion. In the investigation of the representability of an integer as a sum of squares,* however, we had occasion to consider the arithmetic structure of the set of Gaussian integers, and to apply this to a problem involving ordinary integers. During the last century, it has been found that many problems in rational arithmetic are treated most naturally by introducing larger sets of "integers" and deducing, from the structure of the extended system, information about the ordinary integers. Of course, as soon as a mathematician begins to work in a new medium, to use a metaphor from art, he finds interesting questions which have little or nothing to do with the original problem. In the present case, this tendency was instrumental in the development of modern abstract algebra, a large portion of which has only a tenuous connection with number theory.

From the point of view of this text, general algebraic theory must take second place, the primary object being to give the reader an appreciation of the power afforded by the method, as well as a knowledge of some of the basic results in the subject. For this reason, the formulation will be kept as concrete as possible; there will be no striving for generality or abstractness for their own sakes. The treatment is self-contained, except for the following two theorems, whose proofs can be found, for example, in L. E. Dickson, *First Course in the Theory of Equations* (New York: John Wiley & Sons, Inc., 1921), pp. 130–131 and 124–125, respectively.

---

*See, for example, Volume I, Chapter 7.

The product $D_1 D_2$ of two determinants of the same order is another determinant of that order, whose element in row i and column j is the sum of the products of the elements of the ith row of D_1 and the corresponding elements of the jth row of D_2.

SYMMETRIC FUNCTION THEOREM. *Any polynomial* $P(x_1, \ldots, x_n)$ *symmetric in* $x_1, \ldots, x_n$ *and of degree* g *in each, is equal to a polynomial of total degree* g, *with integral coefficients, in the elementary symmetric functions*

$$\sum x_1, \qquad \sum x_1 x_2, \qquad \ldots, \qquad x_1 x_2 \cdots x_n$$

and the coefficients of $P(x_1, \ldots, x_n)$. *In particular, any symmetric polynomial with integral coefficients is equal to a polynomial in the elementary symmetric functions with integral coefficients.*

If P *is a polynomial in the roots of an equation* $f(x) = 0$ *of degree* n *and leading coefficient 1, and if* P *is symmetric in* $n - 1$ *of the roots, then* P *is equal to a polynomial, with integral coefficients, in the remaining root and the coefficients of* $f(x)$ *and* P.

We shall also have occasion to use the so-called Fundamental Theorem of algebra; this basic assertion is proved in the remainder of the section.

FUNDAMENTAL THEOREM OF ALGEBRA. *A polynomial* $f(z) = a_0 z^n + \cdots + a_n$ *having complex coefficients and positive degree, has a complex zero. (It follows immediately that it has exactly* n *complex zeros, in the sense that there are complex numbers* $\xi_1, \ldots, \xi_n$ *such that*

$$f(z) = a_0 (z - \xi_1) \cdots (z - \xi_n).)$$

Proof: Since the truth of the theorem depends on the structure of the complex numbers, it is necessary to use some properties of these numbers. If the entire theory of functions of a complex variable is assumed, the proof is very easy indeed: an analytic function has as many zeros as poles, and a polynomial has a pole at infinity, so it must have at least one zero. If less than this is assumed, it is reasonable to ask that as little be assumed as possible. The proof to be given uses the fact that a real-valued continuous function of two real variables has a minimum value in any closed domain, and it assumes familiarity with the symbol $\sqrt{a}$, where a is real. (If DeMoivre's theorem were used, to give meaning to $\sqrt[n]{a}$ for complex a, the proof would be slightly simpler.)

With the second assumption, the quadratic formula provides a proof when $n = 2$ and the coefficients are real. To solve a quadratic equation with nonreal coefficients, it may be necessary to extract the square root of a nonreal number. Let the number be $a + bi$. Then the equation $a + bi = (x + iy)^2$ gives

$$a = x^2 - y^2 \quad \text{and} \quad b = 2xy,$$

or

$$4x^4 - 4ax^2 - b^2 = 0,$$

and we can take

$$x = \sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}}, \quad y = \frac{b}{2x}.$$

Before treating the general case, note first that we can write $f(x + iy) = G(x, y) + iH(x, y)$, where G and H are polynomials in the real variables x and y, with real coefficients. It follows from the continuity of G and H throughout the xy-plane that $|f(z)|$ is continuous throughout the complex z-plane, where $z = x + iy$. Moreover, for $n > 0$ and $a_0 \neq 0$ (which we henceforth assume), we have

$$\lim_{z \to \infty} |f(z)| = \infty.$$

For if $\max(|a_0|, \ldots, |a_n|) = A$, then

$$|f(z)| \geq |a_0 z^n| - (|a_n| + |a_{n-1}z| + \cdots + |a_1 z^{n-1}|)$$

$$> |a_0 z^n| \left(1 - \frac{nA}{|a_0 z|}\right) \quad \text{for } |z| > 1$$

$$> \frac{|a_0 z^n|}{2} \quad \text{for } |z| > \max\left(\frac{2nA}{|a_0|}, 1\right).$$

Since $|f(z)|$ is continuous, it assumes a minimum value at some point in any closed circular disk with center at O, and since $|f(z)|$ becomes infinite with $|z|$, the disk can be chosen so large that this minimum occurs at an interior point ξ. We must show that $|f(\xi)| = 0$.

We now proceed by induction: suppose that every polynomial of degree less than n, with complex coefficients, has a complex zero, and that f is of degree n and $|f(z)|$ assumes its nonzero minimum at ξ.

Suppose that $f(\xi) = M$, and put

$$g(z) = \frac{f(z + \xi)}{M} = 1 + b_1 z + \cdots + b_n z^n;$$

then $|g(z)| \geq 1$ for all z. Define k as the smallest index such that $b_k \neq 0$, so that

$$g(z) = 1 + b_k z^k + \cdots + b_n z^n, \qquad k \leq n.$$

First consider the case that $k < n$. By the induction hypothesis, the equation

$$1 + b_k z^k = 0$$

has a root. Let η be this root, and put $z = \delta \eta$, where $0 < \delta < 1$. Then

$$g(\delta \eta) = 1 + b_k \delta^k \eta^k + b_{k+1} \delta^{k+1} \eta^{k+1} + \cdots + b_n \delta^n \eta^n$$

$$= 1 - \delta^k + (b_{k+1} \eta^{k+1} + \cdots + b_n \eta^n \delta^{n-k-1}) \delta^{k+1}.$$

Now if $|b_j| < B$ for $k < j \leq n$, then

$$\delta^{k+1} |b_{k+1} \eta^{k+1} + \cdots + b_n \eta^n \delta^{n-k-1}|$$

$$\leq B(1 + |\eta|)^n \delta^{k+1} (1 + \delta + \cdots + \delta^{n-k-1})$$

$$\leq Bn(1 + |\eta|)^n \delta^{k+1} = C \delta^{k+1}.$$

Thus

$$|g(\delta \eta)| < 1 - \delta^k + C \delta^{k+1} = 1 - \delta^k (1 - C\delta),$$

and for $0 < \delta < 1/C$, $|g(\delta \eta)| < 1$. This contradicts the assumption that 1 is the minimum of $|g(z)|$; hence $M = 0$.

If $k = n$, then $g(z) = 1 + b_n z^n$. If n is even, then the equation

$$z^{\frac{1}{2}n} - \sqrt{-\frac{1}{b_n}} = 0$$

is solvable, by the induction hypothesis, and any root of it is also a root of $g(z) = 0$. Hence we can suppose that n is odd. Put $b_n = c + di$. If $c \neq 0$, we put $z = -\delta \operatorname{sgn} c$ (that is, $z = \delta$ or $-\delta$, according as $c < 0$ or $c > 0$), and obtain

$$|1 + (c + di)z^n|^2 = |1 - |c|\delta^n - \delta^n di \operatorname{sgn} c|^2$$

$$= 1 - 2|c|\delta^n + (c^2 + d^2)\delta^{2n};$$

this last expression is again smaller than 1 for δ sufficiently small, and we have the same contradiction as before.

If $c = 0$, then $d \neq 0$; moreover, a sign can be chosen so that $(\pm i)^n = i$. Then if $z = \pm i\delta$ sgn d, we have

$$\left| 1 + id(\pm i\delta \text{ sgn } d)^n \right| = \left| 1 - |d|\delta^n \right|,$$

and this is smaller than 1 for δ sufficiently small. The proof is complete.

2–2 Polynomials and algebraic numbers. We begin by making the following definitions.

(a) R is the set of all rational numbers.

(b) $R[x]$ consists of R together with all polynomials in x with rational coefficients, the coefficient of the highest power of x being different from zero.

(c) If a polynomial $p(x)$ is in $R[x]$, deg p means the exponent of the highest power of x occurring in $p(x)$, if this is positive; if $a \neq 0$ is in R, deg $a = 0$, while if $a = 0$, deg a is not defined.

(d) A polynomial $p(x)$ in $R[x]$ is said to be *monic* if the leading coefficient is 1.

(e) If $p_1(x)$ and $p_2(x)$ are in $R[x]$, we say that $p_2(x)$ *divides* $p_1(x)$ (in symbols, $p_2(x)|p_1(x)$; the phrase *does not divide* is indicated by the symbol "$\nmid$") if there is a $q(x)$ in $R[x]$ such that $p_1(x) = p_2(x)q(x)$. Under this definition, an element of R different from zero divides every element of $R[x]$. The nonzero elements of R are therefore called *units* of $R[x]$.

(f) An element $p(x)$ is said to be *irreducible in* $R[x]$ if it cannot be written as the product of two nonunit elements of $R[x]$.

By formalizing the ordinary process of dividing one polynomial by another, it is not hard to show that if $p_1(x)$ and $p_2(x)$ are in $R[x]$, and $p_2(x)$ is not zero, then there exists a unique pair of elements $q(x)$ and $r(x)$ of $R[x]$ such that

$$p_1(x) = p_2(x)q(x) + r(x), \qquad \deg r < \deg p_2 \text{ or } r(x) = 0.$$

This analog of the division theorem for integers* forms the basis for a Euclidean algorithm, by means of which a greatest common divisor $(p_1(x), p_2(x))$ can be determined; the development is entirely parallel to that for the integers, and leads to the following theorems.

---

*See, for example, Volume I, Theorem 1–1.

THEOREM 2-1. *Given two elements* $p_1(x)$, $p_2(x)$ *of* $R[x]$, *not both zero, there is another element* $d(x)$ *which is unique to within a unit factor and which has the following properties:*

(a) $d(x)|p_1(x)$ *and* $d(x)|p_2(x)$.

(b) *If* $d_1(x)$ *is in* $R[x]$, *and divides both* $p_1(x)$ *and* $p_2(x)$, *then* $d_1(x)|d(x)$.

If $(p_1(x), p_2(x)) = d(x)$, *there are elements* $q_1(x)$ *and* $q_2(x)$ *of* $R[x]$ *such that*

$$p_1(x)q_1(x) + p_2(x)q_2(x) = d(x).$$

THEOREM 2-2. *Any nonzero element of* $R[x]$ *can be factored into a product of irreducible elements of* $R[x]$, *and this factorization is unique except for the order of factors and the presence of units.*

There is no loss in generality, and some gain in simplicity, in supposing that the various polynomials with which we deal are monic, since any polynomial can be made monic by multiplication by a unit. In this case the second part of Theorem 2-2 could be restated to read: *The factorization of a monic polynomial into irreducible monic elements is unique except for the order of factors.*

We now consider the zeros of the polynomials of $R[x]$, or, what is the same thing, the roots of equations $p(x) = 0$. If α is a root of the equation

$$p(x) \equiv x^n + r_1 x^{n-1} + r_2 x^{n-2} + \cdots + r_n = 0. \tag{1}$$

where $p(x)$ is in $R[x]$ and $n > 0$, then α is called an *algebraic number;* if $p(x)$ is irreducible in $R[x]$, α is said to be of *degree n*. (The rational numbers are algebraic numbers, since if r is in R, $x - r = 0$ has the root $x = r$. As algebraic numbers they are of degree 1, although when considered as elements of $R[x]$ the nonzero rational numbers were given degree 0.) An algebraic number α is a zero of a unique monic irreducible polynomial in $R[x]$, called the *defining polynomial* of α. For if $p(x)$ is not irreducible, it can be factored uniquely into irreducible monic factors, and α must be a zero of one of the factors. Hence α satisfies some irreducible equation, i.e., an equation in which the left side is irreducible in $R[x]$. If α satisfies two such equations, say $p(x) = 0$ and $q(x) = 0$, then it also satisfies the equation $d(x) = 0$, where $d(x) = (p(x), q(x))$. For if

$$p(x)s_1(x) + q(x)s_2(x) = d(x),$$

then

$$d(\alpha) = s_1(\alpha) \cdot 0 + s_2(\alpha) \cdot 0 = 0.$$

But since $p(x)$ and $q(x)$ are irreducible, their monic GCD is either 1 or $p(x)$. Since $1 \neq 0$, $(p(x), q(x)) = p(x)$, and $p(x) = q(x)$.

If $p(x)$ in equation (1) is the defining polynomial of α, its n zeros $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$ are called the *conjugates of* α. Except for an alternation in signs, the numbers $r_1, r_2, \ldots, r_n$ are simply the elementary symmetric functions of $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$:

$$r_1 = -\sum \alpha_1 = -(\alpha + \alpha_2 + \cdots + \alpha_n),$$

$$r_2 = \sum \alpha_1 \alpha_2 = \alpha \alpha_2 + \cdots + \alpha_{n-1} \alpha_n,$$

$$\vdots$$

$$r_n = (-1)^n \alpha \alpha_2 \cdots \alpha_n.$$

As is the case here, we shall frequently use a Greek letter, both with and without the subscript 1, to denote a single algebraic number.

THEOREM 2-3. *The sum, difference, and product of two algebraic numbers are algebraic numbers. The quotient of two algebraic numbers is an algebraic number if the denominator is not zero.*

Proof: Suppose that $\alpha = \alpha_1$ and $\beta = \beta_1$ have defining polynomials

$$p(x) = x^n + r_1 x^{n-1} + \cdots + r_n = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_n),$$

$$q(x) = x^m + s_1 x^{m-1} + \cdots + s_m = (x - \beta_1)(x - \beta_2) \cdots (x - \beta_m),$$

respectively. Let $\gamma_1, \gamma_2, \ldots, \gamma_{nm}$ be the numbers obtained by adding an α_i and a β_j, in all possible ways. Then the polynomial $g(x) = (x - \gamma_1)(x - \gamma_2) \cdots (x - \gamma_{nm})$ has, as coefficients, symmetric polynomials in the α_i and β_j, with integral coefficients. Let one such coefficient be $t(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$. As a symmetric polynomial in the α_i it is equal to a polynomial in $r_1, \ldots, r_n$, whose coefficients are themselves polynomials in $\beta_1, \ldots, \beta_m$ with integral coefficients. These last polynomials are symmetric in $\beta_1, \ldots, \beta_m$; they are therefore integral combinations of $s_1, \ldots, s_m$, and consequently are rational numbers. Thus the coefficients of $g(x)$ are rational numbers, and $\alpha + \beta$ is an algebraic number. The same proof applies for $\alpha \cdot \beta$ and $\alpha - \beta$, with obvious changes in the definition of $\gamma_1, \ldots, \gamma_{nm}$.

If α is algebraic and different from zero, so is $1/\alpha$, for the zeros of the polynomial

$$r_n x^n + r_{n-1} x^{n-1} + \cdots + r_1 x + 1$$

are the reciprocals of those of

$$x^n + r_1 x^{n-1} + \cdots + r_n,$$

and $r_n \neq 0$. Thus the assertion that α/β is algebraic is a consequence of the fact that $\alpha \cdot \dfrac{1}{\beta}$ is algebraic.

The properties of the set of all algebraic numbers mentioned in Theorem 2–3 are shared by many sets of importance in mathematics; so many in fact that the name *field* has been reserved to describe such sets. Technically, a field F is a set of two or more elements $a, b, \ldots,$ together with an equivalence relation (which we designate by an equals sign) and two operations (which we designate by the symbols "+" and "·"), such that the following relations hold:

(a) For any a and b in F, either $a = b$ or $a \neq b$. If $a = b$, then $a + c = b + c$ and $a \cdot c = b \cdot c$, for every c in F.

(b) The elements form a commutative group with respect to the operation "+", the identity element being designated by "0". In other words, if a, b, and c are in F, then $a + b$ is in F, $a + b = b + a$, $a + (b + c) = (a + b) + c$, there is an element $-a$ in F such that $a + (-a) = (-a) + a = 0$, and $a + 0 = 0 + a = a$.

(c) The elements with 0 omitted (which we might call F^*) form a commutative group with respect to the operation "·", the identity element being designated by "1".

(d) Multiplication is distributive with respect to addition; that is, $a \cdot (b + c) = a \cdot b + a \cdot c$ for every a, b, and c in F.

As long as one is working with a set of real or complex numbers, and ordinary multiplication, addition, and equality, one can show that the set forms a field just by showing that if a and b are in the set, so are $a \pm b$, ab, and a/b if $b \neq 0$; the other requirements are automatically fulfilled. Thus Theorem 2–3 is just the assertion that the set of all algebraic numbers is a field. Other familiar examples of fields are the set of all rational numbers, the set of all real numbers, and the set of all complex numbers. (The integers, on the other hand, do not form a field, since only the elements ± 1 have inverses, under multiplication, in the system.) In fact, every field composed of

complex numbers together with the ordinary operations of addition and multiplication, contains the field R of rational numbers as a subfield. There are, however, fields with only finitely many elements. An example of such a field is the set of numbers $0, 1, \ldots, p - 1$ with the operations of addition and multiplication modulo p; in this case, $a + b$ is that element c such that $a + b \equiv c \pmod{p}$; $a \cdot b$ is that element d such that $a \cdot b \equiv d \pmod{p}$; $-a$ is 0 or $p - a$, according as a is 0 or not 0; if $a \neq 0$, a^{-1} is that element f such that $a \cdot f \equiv 1 \pmod{p}$.

The field of all algebraic numbers will play no role in the present discussion. We consider instead certain subfields of it, called *algebraic number fields*, described in the next theorem.

Let ϑ be an algebraic number, of degree $n > 1$, whose defining polynomial is $p(x)$ as given in equation (1), and whose conjugates are $\vartheta, \vartheta_2, \ldots, \vartheta_n$.

THEOREM 2-4. *The set of all numbers of the form*

$$\alpha = \frac{q_1(\vartheta)}{q_2(\vartheta)}, \tag{2}$$

where $q_1(x)$ and $q_2(x)$ are in $R[x]$ and $q_2(\vartheta) \neq 0$, is a field, which will be denoted by $R(\vartheta)$. Every element of $R(\vartheta)$ can be expressed uniquely in the form

$$\alpha = a_0 + a_1\vartheta + \cdots + a_{n-1}\vartheta^{n-1},$$

where $a_0, a_1, \ldots, a_{n-1}$ are in R.

Proof: The first part is clear, since the sum, difference, product and quotient of rational functions are again rational functions.

Since $q_2(\vartheta) \neq 0$ and $p(x)$ is irreducible, $q_2(x)$ and $p(x)$ are relatively prime, and for some $t(x)$ and $s(x)$ in $R[x]$,

$$t(x)p(x) + s(x)q_2(x) = 1.$$

This gives $s(\vartheta)q_2(\vartheta) = 1$, and

$$\alpha = \frac{q_1(\vartheta)}{q_2(\vartheta)} = s(\vartheta)q_1(\vartheta),$$

a polynomial in ϑ. Since $p(\vartheta) = 0$,

$$\vartheta^n = -r_1\vartheta^{n-1} - r_2\vartheta^{n-2} - \cdots - r_n.$$

It follows that every positive power of ϑ can be written as a poly-

nomial in ϑ of degree $n-1$ or less. The same is therefore true of every element α. If there were two different representations of α as polynomials in ϑ of degree $n-1$ or less and with rational coefficients, their difference would be a polynomial of degree $n-1$ or less which vanishes for $x = \vartheta$, which is impossible.

If α is an element of the field described in Theorem 2–4, and

$$\alpha = a_0 \vartheta^{n-1} + \cdots + a_{n-1} = \varphi(\vartheta),$$

then the numbers

$$\alpha' = \alpha, \qquad \alpha'' = \varphi(\vartheta_2), \qquad \ldots, \qquad \alpha^{(n)} = \varphi(\vartheta_n)$$

are called the *field conjugates* of α. (They may not lie in the field described in Theorem 2–4.) Every field conjugate of α is also a conjugate of α in the earlier sense, for if α has the defining equation $g(x) = 0$, then $g(\varphi(x))$ vanishes for $x = \vartheta$, so that $p(x)|g(\varphi(x))$ and $g(\varphi(\vartheta_k)) = g(\alpha^{(k)}) = 0$. The converse is also true, as the following theorem shows.

THEOREM 2–5. *The set of field conjugates of an element α of $R(\vartheta)$ is either identical with the set of conjugates of α, or consists of several copies of the set of conjugates of α. (Hence $\deg \alpha | \deg \vartheta$.) The polynomial whose zeros are the field conjugates of α is a power of the defining polynomial of α; if it is equal to the defining polynomial, then $R(\alpha) = R(\vartheta)$.*

Proof: Form the *field polynomial* for α:

$$f(x) = (x - \alpha')(x - \alpha'') \cdots (x - \alpha^{(n)}).$$

Its coefficients are symmetric polynomials in the $\alpha^{(k)}$'s, and are therefore symmetric polynomials in $\vartheta_1, \ldots, \vartheta_n$, and so are rational numbers. Factor $f(x)$ into its monic irreducible factors in $R[x]$, say

$$f(x) = f_1(x) \cdot f_2(x) \cdots,$$

and let $f_1(x)$ be a factor which vanishes for $x = \alpha$. Then $f_1(\varphi(\vartheta)) = 0$, so $p(x)|f_1(\varphi(x))$, and $f_1(x)$ vanishes at $\alpha', \alpha'', \ldots, \alpha^{(n)}$. If these are distinct, $f_1(x)$ is of degree n, and $f(x)$ is irreducible. If they are not, let $\alpha, \alpha'', \ldots, \alpha^{(t)}$ be a maximal distinct set of α's. Then $f_2(x)$ vanishes for some $\alpha^{(k)}$, so $f_1(x)|f_2(x)$; since $f_2(x)$ is irreducible, $f_2(x) = cf_1(x)$, and $c = 1$ since $f_1(x)$ and $f_2(x)$ are monic. If there

are other factors of $f(x)$, the argument can be repeated. Eventually, we find that

$$f(x) = (f_1(x))^{n/t}.$$

Since the zeros of $f_1(x)$, which is the defining polynomial of α, are the conjugates of α, those of $f(x)$ (that is, the field conjugates) consist of n/t copies of the set of conjugates of α.

Now suppose that $f_1(x) \neq f(x)$. Define

$$\varphi(x) = f(x) \left[\frac{\vartheta}{x - \alpha'} + \frac{\vartheta_2}{x - \alpha''} + \cdots + \frac{\vartheta_n}{x - \alpha^{(n)}} \right],$$

so that $\varphi(x)$ is a polynomial of degree $n - 1$ with rational coefficients. Since

$$\varphi(\alpha) = \vartheta(\alpha - \alpha'') \cdots (\alpha - \alpha^{(n)}) = \vartheta f'(\alpha),$$

we have that the number

$$\vartheta = \frac{\varphi(\alpha)}{f'(\alpha)}$$

is in $R(\alpha)$, so that $R(\vartheta)$ is a subfield of $R(\alpha)$, and $R(\alpha) = R(\vartheta)$.

The last assertion of the theorem shows that if one field $R(\alpha)$ is a proper subfield of a second field $R(\vartheta)$, then $\deg \alpha < \deg \vartheta$. For if $\deg \alpha = \deg \vartheta$, then the field polynomial of α with respect to $R(\vartheta)$ is irreducible, so that $f_1(x) = f(x)$, and $R(\alpha) = R(\vartheta)$.

The field $R(\vartheta)$ is called an *algebraic number field;* we say that $R(\vartheta)$ is obtained by *adjoining* ϑ to R, and call $R(\vartheta)$ a *simple algebraic extension* of R, of *degree* n. This same field can be obtained by adjoining various other numbers to R; for example, $R(2\vartheta) = R(\vartheta)$. If an element α of $R(\vartheta)$ is such that $R(\alpha) = R(\vartheta)$, then α is called a *primitive element* of $R(\vartheta)$. It is clear that the degrees of any two primitive elements are the same, and both are equal to the degree of the field.

There is, of course, no reason why the process of adjunction cannot be repeated; one can start from $R(\vartheta)$ and adjoin an algebraic number η to it by taking all rational functions of η whose coefficients are elements of $R(\vartheta)$. This new field is denoted by $R(\vartheta)(\eta)$, or more simply by $R(\vartheta, \eta)$.

THEOREM 2–6. *If ϑ and η are algebraic numbers, the adjunction of η to $R(\vartheta)$ gives the same field $R(\vartheta, \eta)$ as the adjunction of ϑ to $R(\eta)$.*

There exists an algebraic number ζ such that $R(\vartheta, \eta)$ is identical with $R(\zeta)$.

Proof: The first part is clear, since both $R(\vartheta, \eta)$ and $R(\eta, \vartheta)$ are identical with the field consisting of all numbers of the form

$$\frac{q_1(\vartheta, \eta)}{q_2(\vartheta, \eta)}, \qquad q_2(\vartheta, \eta) \neq 0,$$

where $q_1(x, y)$ and $q_2(x, y)$ are polynomials in two variables with rational coefficients.

If η is an element of $R(\vartheta)$, then $R(\vartheta, \eta) = R(\vartheta)$, since a rational function of a rational function is again a rational function. Assume then that ϑ and η do not lie in the fields $R(\eta)$ and $R(\vartheta)$ respectively. Let their defining polynomials be $p_1(x)$ and $p_2(x)$, and let their conjugates be $\vartheta_1, \ldots, \vartheta_n$ and $\eta_1, \ldots, \eta_m$, respectively. Let a and b be rational numbers, and let $\zeta = \zeta_1, \ldots, \zeta_{nm}$ be all expressions of the form $a\vartheta_j + b\eta_k$. Since the conjugates of ϑ are distinct, as are the conjugates of η, there is only a finite set of ratios a/b for which some two of the ζ's are equal, and we choose a and b so that a/b is not in this set. Furthermore, we order the ζ_i so that $\zeta = a\vartheta + b\eta$.

Now put

$$f(x) = (x - \zeta_1)(x - \zeta_2) \cdots (x - \zeta_{nm}).$$

This polynomial has no multiple zeros, and its coefficients, being symmetric in the ϑ's and η's, are rational. We show that $R(\vartheta, \eta) = R(\zeta)$. It is clear that every element of $R(\zeta)$ is in $R(\vartheta, \eta)$. Suppose on the other hand that ρ is in $R(\vartheta, \eta)$, and that

$$\rho = \frac{q_1(\vartheta, \eta)}{q_2(\vartheta, \eta)}, \qquad q_2(\vartheta, \eta) \neq 0.$$

Then we can define the numbers $\rho = \rho_1, \ldots, \rho_{nm}$ by the equation

$$\rho_i = \frac{q_1(\vartheta_j, \eta_k)}{q_2(\vartheta_j, \eta_k)},$$

where the same subscripts appear on ϑ and η in the definition of ρ_i as in the definition of ζ_i, for $i = 1, 2, \ldots, nm$. Now put

$$F(x) = f(x)\left(\frac{\rho}{x - \zeta_1} + \frac{\rho_2}{x - \zeta_2} + \cdots + \frac{\rho_{nm}}{x - \zeta_{nm}}\right);$$

by the Symmetric Function Theorem, the coefficients of $F(x)$ are rational. If $i > 1$, the polynomial

$$f(x) \frac{\rho_i}{x - \zeta_i} = \rho_i(x - \zeta) \cdots (x - \zeta_{i-1})(x - \zeta_{i+1}) \cdots (x - \zeta_{nm})$$

vanishes for $x = \zeta$, and from the representation

$$f(x) \frac{\rho}{x - \zeta} = \rho(x - \zeta_2) \cdots (x - \zeta_{nm})$$

we have

$$F(\zeta) = \rho(\zeta - \zeta_2) \cdots (\zeta - \zeta_{nm}).$$

Since

$$f'(\zeta) = (\zeta - \zeta_2) \cdots (\zeta - \zeta_{nm}) \neq 0,$$

this gives

$$\rho = \frac{F(\zeta)}{f'(\zeta)},$$

and ρ is in $R(\zeta)$.

PROBLEMS

1. Prove *Eisenstein's irreducibility criterion:* a polynomial $f(x) = a_0 + a_1x + \cdots + a_nx^n$ with integral coefficients cannot be written as a product of two or more polynomials with integral coefficients and positive degrees, if there is a prime p such that

$$p \nmid a_n, \qquad p \mid a_i \text{ if } i < n, \qquad \text{and } p^2 \nmid a_0.$$

[*Hint:* Suppose that there is such a p, but that $f(x) = g(x)h(x)$, where $g(x) = b_0 + b_1x + \cdots + b_rx^r$, $h(x) = c_0 + c_1x + \cdots + c_sx^s$. It follows that p divides exactly one of b_0 and c_0—say b_0. Let b_i be the first coefficient in $g(x)$ not divisible by p, and deduce a contradiction from the expression for a_i in terms of the b's and c's.] As we shall see later (Theorem 2–21), irreducibility over the set of polynomials with integral coefficients implies irreducibility over $R[x]$. Use this fact in Problem 2.

2. Show that the following polynomials are irreducible over $R[x]$:

    (a) $x^n - p$, p a prime.

    (b) $x^{p-1} + x^{p-2} + \cdots + x + 1$. [*Hint:* Replace x by $x + 1$.]

    (c) $x^3 + 3x^2 + 4$.

3. Show that $R(\sqrt{2}, \sqrt{3})$ is identical with $R(\sqrt{2} + \sqrt{3})$, and find a rational function $r(x)$ with rational coefficients such that $r(\sqrt{2} + \sqrt{3}) = \sqrt{2}$.

2–3 Algebraic integers. If the defining (monic) polynomial of an algebraic number ϑ has integral coefficients, ϑ is said to be an *algebraic integer*. This is a direct extension of the notion of ordinary or *rational* integers, which are the zeros of monic linear polynomials with integral coefficients. Hereafter we shall designate by Z the set of all rational integers.

THEOREM 2–7. *The sum, difference, and product of two algebraic integers are again algebraic integers.*

The proof follows the lines of the proof of Theorem 2–3.

THEOREM 2–8. *If α is a zero of a monic polynomial with coefficients in Z, then α is an algebraic integer.*

Proof: Suppose that $f(x) = x^N + \cdots + a_N$ is the polynomial, and that $p(x) = x^n + r_1 x^{n-1} + \cdots + r_n$ is the defining polynomial of . Let b_0 be the LCM of the denominators of the reduced fractions $r_1, \ldots, r_n$, so that $b_0 p(x) = q(x) = b_0 x^n + b_1 x^{n-1} + \cdots + b_n$ has relatively prime rational integral coefficients. Then $q(x)$ divides $f(x)$, the coefficients in the quotient polynomial being rational, and we can write

$$\frac{f(x)}{q(x)} = \frac{cg(x)}{c'},$$

where c and c' are so chosen that $g(x)$ has relatively prime coefficients in Z. Thus $c'f(x) = cg(x)q(x)$, and the coefficients of the product $g(x)q(x)$ are relatively prime.* Since this is also true of the coefficients of $f(x)$ (for $f(x)$ is monic), we conclude that $c = c'$. Comparing the coefficients of x^N in the equation $f(x) = g(x)q(x)$, we see that $b_0 | 1$, and hence $b_0 = \pm 1$, which was to be proved.

THEOREM 2–9. *If α is a root of an equation*

$$f(x) = x^n + \beta_1 x^{n-1} + \cdots + \beta_n = 0,$$

in which $\beta_1, \ldots, \beta_n$ are algebraic integers, then α is an algebraic integer.

Proof: By Theorem 2–6, $\beta_1, \ldots, \beta_n$ all lie in a simple extension field $R(\vartheta)$, of degree m, say. We can use the sets of field conjugates

_____

*The reader may prove this simple fact himself, or refer to the remark preceding Theorem 3–14, Volume I.

$$(\beta_1{}'', \ldots, \beta_n{}''), \qquad \ldots, \qquad (\beta_1{}^{(m)}, \ldots, \beta_n{}^{(m)})$$

to form polynomials

$$f_2(x) = x^n + \beta_1{}'' x^{n-1} + \cdots + \beta_n{}'', \qquad \ldots,$$
$$f_m(x) = x^n + \beta_1{}^{(m)} x^{n-1} + \cdots + \beta_n{}^{(m)}.$$

The product $f(x)f_2(x) \cdots f_m(x)$ has rational integral coefficients and is monic; by Theorem 2–8, α is an algebraic integer.

The set of integers in a fixed algebraic number field $R(\vartheta)$ is also closed under addition, subtraction, and multiplication. We shall designate this set by $R[\vartheta]$, and call it the *integral domain* of the field. In particular, $R[1] = Z$ is the set of rational integers.

THEOREM 2–10. *If ϑ is an algebraic number, there exists some rational integer $a \neq 0$ such that $a\vartheta$ is an algebraic integer. If ϑ satisfies an equation $\beta_0 x^n + \cdots + \beta_n = 0$, in which $\beta_0, \ldots, \beta_n$ are algebraic integers, then $\beta_0 \vartheta$ is an algebraic integer.*

Proof: Let the defining equation of ϑ be

$$p(x) = x^n + r_1 x^{n-1} + \cdots + r_n = 0,$$

and let the LCM of the denominators of the fractions $r_1, \ldots, r_n$ be a. Then the polynomial

$$a^n p\left(\frac{x}{a}\right) = x^n + a r_1 x^{n-1} + \cdots + a^n r_n$$

has integral coefficients and is monic and irreducible; its zeros $a\vartheta$, $a\vartheta_2, \ldots, a\vartheta_n$ are therefore integers. The proof of the second part, using Theorem 2–9, is similar.

Since $R(\vartheta)$ and $R(a\vartheta)$ are identical for $a \neq 0$ in Z, any algebraic number field can be considered as the result of adjoining an algebraic integer to R.

If ϑ is an integer, so are its conjugates $\vartheta_2, \ldots, \vartheta_n$. The same is therefore true of its field conjugates.

If α is any element of the field $R(\vartheta)$ of degree n, the product $\alpha \alpha'' \cdots \alpha^{(n)}$ of all the field conjugates of α is called the *norm* of α, and denoted by $\mathbf{N}\alpha$ (a more complete notation would be $\mathbf{N}_{R(\vartheta)}\alpha$).

THEOREM 2–11. *The norm of an algebraic integer is a rational integer.*

Proof: If α has the defining equation

$$x^m + s_1 x^{m-1} + \cdots + s_m,$$

then the norm of α (in any given $R(\vartheta)$ containing α) is a power of s_m, by Theorem 2-5.

THEOREM 2-12. *If α and β are elements of $R(\vartheta)$, then*

$$\mathbf{N}(\alpha\beta) = \mathbf{N}\alpha \cdot \mathbf{N}\beta.$$

Proof: Put

$$\alpha = a_0 + a_1\vartheta + \cdots + a_{n-1}\vartheta^{n-1},$$
$$\beta = b_0 + b_1\vartheta + \cdots + b_{n-1}\vartheta^{n-1}. \tag{3}$$

Then in the product $\alpha\beta$, powers of ϑ higher than the $(n-1)$th can be reduced using the equation

$$\vartheta^{n+j} = -\vartheta^j(r_1\vartheta^{n-1} + \cdots + r_n) \tag{4}$$

derived from the defining equation of ϑ. Also $\alpha^{(k)}$ and $\beta^{(k)}$ can be obtained from (3) by replacing ϑ by ϑ_k, and in the product $\alpha^{(k)}\beta^{(k)}$, higher powers of ϑ_k can be reduced by using (4) with ϑ replaced by ϑ_k. Hence the field conjugates $(\alpha\beta)'$, $(\alpha\beta)''$, ..., $(\alpha\beta)^{(n)}$ of $\alpha\beta$ are simply $\alpha\beta$, $\alpha''\beta''$, ..., $\alpha^{(n)}\beta^{(n)}$. Thus

$$\mathbf{N}\alpha\beta = (\alpha\beta)'(\alpha\beta)'' \cdots (\alpha\beta)^{(n)}$$
$$= \alpha'\alpha'' \cdots \alpha^{(n)}\beta'\beta'' \cdots \beta^{(n)} = \mathbf{N}\alpha \cdot \mathbf{N}\beta.$$

Now let α, β, ..., ν be n elements of $R(\vartheta)$, with field conjugates $\alpha^{(k)}$, $\beta^{(k)}$, ..., $\nu^{(k)}$, where $k = 1, 2, \ldots, n$. The number

$$\Delta(\alpha, \beta, \ldots, \nu) = \begin{vmatrix} \alpha & \alpha'' & \cdots & \alpha^{(n)} \\ \beta & \beta'' & \cdots & \beta^{(n)} \\ \vdots & \vdots & & \vdots \\ \nu & \nu'' & \cdots & \nu^{(n)} \end{vmatrix}^2$$

is called the *discriminant* of α, β, ..., ν. Its value is independent of the order of rows or of columns.

THEOREM 2-13. *If α, β, ..., ν are in $R[\vartheta]$, then $\Delta(\alpha, \beta, \ldots, \nu)$ is a rational integer.*

Proof: If we take the row-by-column product, we have

$$
\Delta(\alpha, \ldots, \nu) =
\begin{vmatrix}
\alpha & \cdots & \alpha^{(n)} \\
\vdots & & \vdots \\
\nu & \cdots & \nu^{(n)}
\end{vmatrix}
\cdot
\begin{vmatrix}
\alpha & \cdots & \nu \\
\vdots & & \vdots \\
\alpha^{(n)} & \cdots & \nu^{(n)}
\end{vmatrix}
$$

$$
=
\begin{vmatrix}
\alpha^2 + \cdots + (\alpha^{(n)})^2 & \cdots & \alpha\nu + \cdots + \alpha^{(n)}\nu^{(n)} \\
\vdots & & \vdots \\
\alpha\nu + \cdots + \alpha^{(n)}\nu^{(n)} & \cdots & \nu^2 + \cdots + (\nu^{(n)2})
\end{vmatrix}
.
$$

Just as in the proof of Theorem 2–12,

$$
\alpha\beta + \alpha''\beta'' + \cdots + \alpha^{(n)}\beta^{(n)} = \alpha\beta + (\alpha\beta)'' + \cdots + (\alpha\beta)^{(n)},
$$

and the sum of the field conjugates of an integer is itself a rational integer, by analogy with the proof of Theorem 2–11. Hence, the number $\Delta(\alpha, \beta, \ldots, \nu)$ can be written as a determinant with rational integral entries, and so is a rational integer.

The numbers $1, \vartheta, \ldots, \vartheta^{n-1}$ are said to form a *basis* of $R(\vartheta)$, in the sense that every element of $R(\vartheta)$ can be expressed in a unique way as a linear combination of these numbers, with coefficients in R (cf. Theorem 2–4). We now examine the possibility of finding a basis for $R[\vartheta]$; that is, a set of elements of $R[\vartheta]$ such that every element of $R[\vartheta]$ can be expressed in a unique way as a linear combination of them, the coefficients in this case being in Z. To emphasize the distinction between these two kinds of bases, the second is sometimes called an *integral basis*. Every integral basis is a basis of $R(\vartheta)$, as is immediately seen from Theorem 2–10, but the converse is false.

If $\omega_1, \ldots, \omega_n$ is to be an integral basis, then for any ρ in $R[\vartheta]$, the equation

$$
\rho = x_1\omega_1 + \cdots + x_n\omega_n,
$$

and therefore also the equations

$$
\rho^{(k)} = x_1\omega_1^{(k)} + \cdots + x_n\omega_n^{(k)}, \qquad k = 2, \ldots, n
$$

must hold for some rational integers $x_1, \ldots, x_n$. If $\Delta(\omega_1, \ldots, \omega_n) \neq 0$, this system of equations can be solved, giving each x_i as the quotient of determinants, the determinant in each denominator being a square root of $\Delta(\omega_1, \ldots, \omega_n)$. It seems plausible that the smaller $|\Delta(\omega_1, \ldots, \omega_n)|$, the better the chance of obtaining rational integral

x_i. Hence, if an integral basis always exists, the next theorem ought to be true.

THEOREM 2–14. *If $\omega_1, \omega_2, \ldots, \omega_n$ are any n integers of $R[\vartheta]$ for which $|\Delta(\omega_1, \omega_2, \ldots, \omega_n)|$ has its smallest possible value different from zero, then $\omega_1, \ldots, \omega_n$ form a basis of $R[\vartheta]$.*

Proof: Write

$$\omega_i = \sum_{j=0}^{n-1} a_{ij}\vartheta^j, \qquad i = 1, 2, \ldots, n \tag{5}$$

where the a_{ij} are in R. Then

$$\Delta(\omega_1, \ldots, \omega_n) = \begin{vmatrix} \omega_1 & \cdots & \omega_n \\ \vdots & & \vdots \\ \omega_1^{(n)} & \cdots & \omega_n^{(n)} \end{vmatrix}^2 = \begin{vmatrix} \sum_{j=0}^{n-1} a_{1j}\vartheta^j & \cdots & \sum_{j=0}^{n-1} a_{nj}\vartheta^j \\ \vdots & & \vdots \\ \sum_{j=0}^{n-1} a_{1j}\vartheta_n^j & \cdots & \sum_{j=0}^{n-1} a_{nj}\vartheta_n^j \end{vmatrix}^2,$$

and this can be factored:

$$\Delta(\omega_1, \ldots, \omega_n) = \left\{ \begin{vmatrix} 1 & \vartheta & \cdots & \vartheta^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \vartheta_n & \cdots & \vartheta_n^{n-1} \end{vmatrix} \cdot \begin{vmatrix} a_{10} & \cdots & a_{n0} \\ \vdots & & \vdots \\ a_{1,n-1} & \cdots & a_{n,n-1} \end{vmatrix} \right\}^2$$

$$= (\det|a_{ij}|)^2 \Delta(1, \vartheta, \ldots, \vartheta^{n-1}), \tag{6}$$

where $\vartheta, \vartheta_2, \ldots, \vartheta_n$ are the conjugates of ϑ. Since $\Delta(\omega_1, \ldots, \omega_n) \neq 0$, also $\det|a_{ij}| \neq 0$, and the system of equations (5) can be solved for the numbers $1, \vartheta, \ldots, \vartheta^{n-1}$, giving linear expressions in $\omega_1, \ldots, \omega_n$. Thus every number ρ of $R[\vartheta]$ can be written in the form

$$\rho = b_1\omega_1 + \cdots + b_n\omega_n, \tag{7}$$

where $b_1, \ldots, b_n$ are rational. We must show that they are rational integers.

If this is not the case for the ρ of (7), then some b_i has a nonzero fractional part:

$$b_i = [b_i] + c,$$

where $0 < c < 1$, and the symbol $[b]$ means the largest integer not exceeding b. Put

$$\rho_i = \rho - [b_i]\omega_i = b_1\omega_1 + \cdots + c\omega_i + \cdots + b_n\omega_n.$$

In just the same way that (6) was deduced from (5), we can deduce
from the system of equations

$$\omega_1 = \omega_1,$$
$$\omega_2 = \omega_2,$$
$$\vdots$$
$$\omega_{i-1} = \omega_{i-1},$$
$$\rho_i = b_1\omega_1 + b_2\omega_2 + \cdots + c\omega_i + \cdots + b_n\omega_n,$$
$$\omega_{i+1} = \omega_{i+1},$$
$$\vdots$$
$$\omega_n = \omega_n,$$

the relation

$$\Delta(\omega_1, \ldots, \rho_i, \ldots, \omega_n) = \begin{vmatrix} 1 & 0 & \ldots & & 0 \\ 0 & 1 & \ldots & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ b_1 & b_2 & \ldots c \ldots & & b_n \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & & 1 \end{vmatrix}^2 \Delta(\omega_1, \omega_2, \ldots, \omega_n)$$

$$= c^2 \Delta(\omega_1, \ldots, \omega_n).$$

But this implies that the discriminant of the system $\omega_1, \ldots, \rho_i, \ldots, \omega_n$
is numerically smaller than that of $\omega_1, \ldots, \omega_n$, and is not zero, which
is contrary to the hypothesis that $|\Delta(\omega_1, \ldots, \omega_n)|$ is minimal.

Any two integral bases of a single field have the same discriminant,
since each is the product of the other and the square of a determinant
with integral entries, as in (6). The common value is called the
discriminant of the field; we shall designate it by Δ hereafter.

PROBLEMS

1. Let $\vartheta, \vartheta',$ and ϑ'' be the roots of
   (a) $x^3 + 2x + 6 = 0,$
   (b) $x^3 - x^2 - x - 2 = 0.$
   Compute the numbers $N_{R(\vartheta)}(3\vartheta - 2).$
   Answer: (a) -206; (b) $4, 19$.

2. (a) Let $f(x) = a_0 x^n + \cdots + a_n$ be irreducible over R, and let $\vartheta, \vartheta'', \ldots, \vartheta^{(n)}$ be the zeros of f. Show that in $R(\vartheta)$,

$$a_0{}^n \Delta(1, \vartheta, \ldots, \vartheta^{n-1}) = (-1)^{n(n-1)/2} \prod_{i=1}^n f'(\vartheta^{(i)}).$$

[This depends on the well-known factorization

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ x_1{}^{n-1} & x_2{}^{n-1} & \cdots & x_n{}^{n-1} \end{vmatrix} = \prod_{1 \le j < i \le n} (x_i - x_j)$$

of a Vandermonde determinant.]

(b) If in particular $f(x) = x^3 + px + q$, show that $\Delta(1, \vartheta, \vartheta^2) = -27q^2 - 4p^3$.

3. Show that if $\alpha_1, \ldots, \alpha_n$ are elements of $R[\vartheta]$ such that $\Delta(\alpha_1, \ldots, \alpha_n)$ is square-free, then $\alpha_1, \ldots, \alpha_n$ form a basis for $R[\vartheta]$.

2–4 Units and primes in $R[\vartheta]$. If α and β are in an integral domain $R[\vartheta]$, we say that β *divides* α, and write $\beta|\alpha$, if there is another element γ of $R[\vartheta]$ such that $\alpha = \beta\gamma$. An integer ϵ such that $\epsilon|1$ is called a *unit* of $R[\vartheta]$. We say that α and β are *associates* if $\alpha = \epsilon\beta$, where ϵ is a unit.

THEOREM 2–15. *An element of $R[\vartheta]$ is a unit if and only if its norm (as an element of $R(\vartheta)$) is ± 1.*

Proof: If ϵ is a unit and

$$x^m + e_1 x^{m-1} + \cdots + e_m = 0$$

is its defining equation, then the defining equation of $1/\epsilon$ is

$$x^m + \frac{e_{m-1}}{e_m} x^{m-1} + \cdots + \frac{1}{e_m} = 0.$$

Since $1/\epsilon$ is an integer, $e_m = \pm 1$, and $\mathbf{N}(1/\epsilon)$ is a power of the constant term in the defining equation of $1/\epsilon$. (Alternatively, this result could be deduced from the multiplicativity of the norm. For if ϵ is a unit, there exists an integer ϵ_1 such that $\epsilon\epsilon_1 = 1$. Hence $1 = \mathbf{N}1 = \mathbf{N}\epsilon\epsilon_1 = \mathbf{N}\epsilon \cdot \mathbf{N}\epsilon_1$, and since the norm of an integer is a rational integer, $\mathbf{N}\epsilon = \pm 1$.)

Conversely, if the constant term in the defining equation of an element of $R[\vartheta]$ is ± 1, then the reciprocal of the element is also an element of $R[\vartheta]$, and the element is a unit.

The units of an integral domain form a multiplicative group, since the product of units is a unit, 1 is a unit, and each unit ϵ has an inverse ϵ_1 such that $\epsilon\epsilon_1 = 1$.

In the domain of rational integers, the only units are ± 1; in the Gaussian domain $R[i]$, the units are ± 1, $\pm i$. All these units are roots of unity, but in some domains there are units which are not roots of unity, and in fact do not have absolute value 1. This was pointed out in Chapter 8 of Volume I, but we can now go into details.

Let d be a square-free rational integer, and consider the field $R(\sqrt{d})$. As a basis for the field we can take 1, $\sqrt{d}$, so that every element of $R(\sqrt{d})$ can be uniquely expressed in the form $a + b\sqrt{d}$, where a and b are in R. If $b = 0$, then $a + b\sqrt{d}$ is an integer if and only if a is in Z. If $b \neq 0$, the defining equation of $a + b\sqrt{d}$ is

$$(x - a - b\sqrt{d})(x - a + b\sqrt{d}) = x^2 - 2ax + a^2 - db^2 = 0,$$

so that if $a + b\sqrt{d}$ is in $R[\sqrt{d}]$, both $2a$ and $a^2 - db^2$ must be rational integers. Hence $(2a)^2 - 4(a^2 - db^2) = 4db^2$ is also in Z; since d is square-free, it follows that $2b$ is in Z.

Suppose that $a = k + \frac{1}{2}$, with k in Z. Then

$$0 \equiv 4a^2 - 4db^2 \equiv 4k^2 + 4k + 1 - 4db^2 \equiv 1 - 4db^2 \pmod 4,$$

and it follows that $2b \equiv 1 \pmod 2$, and $d \equiv 1 \pmod 4$. Conversely, if a and b are halves of odd integers and $d \equiv 1 \pmod 4$, the defining equation of $a + b\sqrt{d}$ has coefficients in Z. Hence 1 and $(1 + \sqrt{d})/2$ form a basis of $R[\sqrt{d}]$, if $d \equiv 1 \pmod 4$.

If $d \equiv 2$ or $3 \pmod 4$, then a must be a rational integer. If b were of the form $k + \frac{1}{2}$, with k in Z, we should have

$$0 \equiv 4a^2 - 4db^2 \equiv -(4k^2 + 4k + 1)d \equiv -d \pmod 4,$$

and d would not be square-free. Hence in this case both a and b must be in Z, and 1, $\sqrt{d}$ form a basis of $R[\sqrt{d}]$.

THEOREM 2–16. *Let d be a square-free rational integer. Then if $d \equiv 1 \pmod 4$, the elements of $R[\sqrt{d}]$ are either of the form*

$$a + b\sqrt{d}, \quad a \text{ and } b \text{ in } Z, \tag{8}$$

or

$$\frac{a + b\sqrt{d}}{2}, \quad a \text{ and } b \text{ in } Z, \quad a \equiv b \equiv 1 \pmod 2,$$

and the discriminant of $R(\sqrt{d})$ is

$$\Delta = \begin{vmatrix} 1 & \frac{1}{2}(1 + \sqrt{d}) \\ 1 & \frac{1}{2}(1 - \sqrt{d}) \end{vmatrix}^2 = (-\sqrt{d})^2 = d.$$

If $d \equiv 2$ or 3 (mod 4), all the elements of $R[\sqrt{d}]$ are of the form (8), and the discriminant of $R(\sqrt{d})$ is

$$\Delta = \begin{vmatrix} 1 & \sqrt{d} \\ 1 & -\sqrt{d} \end{vmatrix}^2 = (-2\sqrt{d})^2 = 4d.$$

The units of $R[\sqrt{d}]$ are the integers ϵ for which $\mathbf{N}\epsilon = \pm 1$. If $d \equiv 2$ or 3 (mod 4), then ϵ is of the form (8), so that the units are given by the solutions of the Pell equations

$$x^2 - dy^2 = \pm 1. \tag{9}$$

If $d \equiv 1$ (mod 4), the units are the integers of the form $(x + y\sqrt{d})/2$, where $x + y\sqrt{d}$ is a solution of one of the Pell equations

$$x^2 - dy^2 = \pm 4. \tag{10}$$

If $d < 0$, these Pell equations have only trivial solutions: (9) has solutions ± 1, 0 in all cases, and 0, ± 1 if $d = -1$, while (10) has the solution ± 2, 0 always, and ± 1, ± 1 if $d = -3$. If $d > 0$, equations (9) and (10) have infinitely many solutions.*

Returning to the general domain $R[\vartheta]$, we say that an element π is *prime* if it is not a unit and has no factors other than its associates and units.

THEOREM 2–17. *Every nonunit element of $R[\vartheta]$ can be written as a finite product of primes.*

Proof: If α in $R[\vartheta]$ is not a unit, $|\mathbf{N}\alpha| > 1$. If α is prime, we have the trivial representation $\alpha = \alpha$. If not, there is a factorization $\alpha = \beta\gamma$ into nonunits, and $\mathbf{N}\alpha = \mathbf{N}\beta \cdot \mathbf{N}\gamma$, where

$$1 < |\mathbf{N}\beta| < |\mathbf{N}\alpha|, \qquad 1 < |\mathbf{N}\gamma| < |\mathbf{N}\alpha|.$$

If either β or γ is not prime, it may be factored. The process must terminate, since the rational integer $\mathbf{N}\alpha$ has only finitely many divisors of absolute value greater than 1.

---

*This result is given in Chapter 8, Volume I. The solutions for given d can be found explicitly with the aid of Theorem 9–6 of that volume.

To see that this factorization need not be unique, consider the two representations

$$21 = 3 \cdot 7 = (4 + \sqrt{-5})(4 - \sqrt{-5})$$

of 21 in $R[\sqrt{-5}]$. Since $-5 \not\equiv 1 \pmod 4$, the integers of this domain are $a + b\sqrt{-5}$, with a and b in Z, and the units are ± 1. It is clear that no two of the numbers $3, 7, 4 + \sqrt{-5}, 4 - \sqrt{-5}$ are associates, and we can also show that all of them are primes in $R[\sqrt{-5}]$. Suppose that

$$(a_1 + b_1\sqrt{-5})(a_2 + b_2\sqrt{-5}) = 3.$$

Then

$$\mathbf{N}(a_1 + b_1\sqrt{-5})\mathbf{N}(a_2 + b_2\sqrt{-5}) = \mathbf{N}3 = 9,$$

so that if neither factor is a unit, it must be that

$$\mathbf{N}(a_1 + b_1\sqrt{-5}) = a_1{}^2 + 5b_1{}^2 = 3. \tag{11}$$

This equation, however, has no solution in Z. By a similar argument, 7 has no proper divisors, since the equation

$$a_1{}^2 + 5b_1{}^2 = 7 \tag{12}$$

has no solution in Z. Finally, an assumed factorization of either $4 \pm \sqrt{-5}$ leads to the equation

$$\mathbf{N}(a_1 + b_1\sqrt{-5}) \cdot \mathbf{N}(a_2 + b_2\sqrt{-5}) = 21,$$

which in turn requires that either (11) or (12) hold. Hence $R[\sqrt{-5}]$ is not a unique factorization domain.

A domain $R[\vartheta]$ is called a *Euclidean domain* if for any pair of integers $\beta \neq 0$ and α of $R[\vartheta]$, there is an element γ such that

$$|\mathbf{N}(\alpha - \beta\gamma)| < |\mathbf{N}\beta|.$$

In this case, there is a Euclidean algorithm by means of which a greatest common divisor can be defined, such that if $(\alpha, \beta) = \delta$, there are integers γ_1 and γ_2 in $R[\vartheta]$ for which $\alpha\gamma_1 + \beta\gamma_2 = \delta$. It is this last property which is essential for unique factorization, since from it we get the result, equivalent to the Unique Factorization Theorem, that if $\beta | \alpha\gamma$ and $(\beta, \alpha) = 1$ then $\beta | \gamma$. For if $\gamma_1\alpha + \gamma_2\beta = 1$, then $\gamma_1\alpha\gamma + \gamma_2\beta\gamma = \gamma$; hence $\beta | \gamma$. There is no such GCD in $R[\sqrt{-5}]$.

For example, 3 and $4 + \sqrt{-5}$ must be considered as relatively prime, since they are nonassociated primes, but if we had

$$3(a + b\sqrt{-5}) + (4 + \sqrt{-5})(c + d\sqrt{-5}) = 1, \qquad a, b, c, d \text{ in } Z$$

it would follow that

$$3a + 4c - 5d = 1, \qquad 3b + c + 4d = 0.$$

Subtracting the second equation from the first, we would have

$$3(a - b + c - 3d) = 1,$$

which is palpably false.

Every Euclidean domain, then, is a unique factorization domain, although the converse is not true. The quadratic Euclidean domains are completely known: $R[\sqrt{d}]$ is Euclidean if and only if d has one of the 21 values $-11, -7, -3, -2, -1, 2, 3, 5, 6, 7, 11, 13, 17, 19, 21, 29, 33, 37, 41, 57,$ or 73.

PROBLEMS

1. Show that $R[\rho]$, where $\rho = (-1 + i\sqrt{3})/2$ is a cube root of unity, is a Euclidean domain. [Compare Theorem 7-6, Volume I.]

2. Find the GCD of $2 + \rho$ and $5 + 7\rho$ in $R[\rho]$.

3. Show that if d is square-free, and if Δ is the discriminant of $R(\sqrt{d})$, then the numbers 1 and $(\Delta + \sqrt{\Delta})/2$ form a basis of $R[\sqrt{p}]$.

2-5 Ideals. One way of restoring unique factorization consists in enlarging the set of possible divisors; we might for example try to find entities $A, B, C,$ and D of $R[\sqrt{-5}]$ which are in some sense prime, and such that

$$3 = AB, \qquad 7 = CD, \qquad 4 + \sqrt{-5} = AC, \qquad 4 - \sqrt{-5} = BD.$$

Then the two representations of 21 in $R[\sqrt{-5}]$ would no longer differ essentially; instead we would have

$$21 = (AB)(CD) = (AC)(BD) = ABCD.$$

To accomplish this without going outside the domain, we make a shift of emphasis; rather than asking for the divisors of a given number, we look for all the numbers which have a given divisor. Here two

properties of the divisibility concept, in which the divisor is fixed, come to mind:

(a) If $\gamma|\alpha$, then $\gamma|\alpha\lambda$ for every integer λ.

(b) If $\gamma|\alpha$ and $\gamma|\beta$, then $\gamma|\alpha \pm \beta$.

In other words, the set of multiples of γ forms an additive group which is closed under multiplication by elements of the domain (but not necessarily in the set). If $\alpha|\beta$, then the set of multiples of α contains the set of multiples of β. The GCD (if there is one) of α and β has as multiples the set of numbers of the form $\alpha' + \beta'$, where α' and β' run independently over the multiples of α and β respectively, and this set is again an additive group closed under multiplication by elements of the domain.

Because of the repeated occurrence of this special kind of set, we give the name *ideal* to any subset (containing at least one element besides zero) of an integral domain $R[\vartheta]$ which forms a group under addition and is closed under multiplication by elements of the domain. Since there is no reason to suppose that every ideal of $R[\vartheta]$ consists of all the multiples of a single element of $R[\vartheta]$, we shall designate a general ideal by a capital letter. A *principal* ideal, consisting of all multiples of a given element α of the domain, will be designated by $[\alpha]$. (It will be clear from the context whether the brackets designate an ideal or the greatest-integer function.) But instead of a single number α, we could begin with any finite set $\alpha_1, \ldots, \alpha_m$ of elements of $R[\vartheta]$, and form all expressions

$$\lambda_1\alpha_1 + \lambda_2\alpha_2 + \cdots + \lambda_m\alpha_m,$$

where $\lambda_1, \ldots, \lambda_m$ run independently over $R[\vartheta]$; the set of such expressions again forms an ideal, which will be designated by $[\alpha_1, \ldots, \alpha_m]$. (The numbers $\alpha_1, \ldots, \alpha_m$ are called *generators* of the ideal $[\alpha_1, \ldots, \alpha_m]$.) This notation is similar to that for the GCD, if such exists, except that instead of writing $(\alpha, \beta) = \gamma$ we would now write $[\alpha, \beta] = [\gamma]$. (Two ideals are said to be equal if they consist of the same numbers.) It will be shown later that $R[\vartheta]$ is a unique factorization domain if and only if every ideal of $R[\vartheta]$ is a principal ideal. This should not be surprising, since this latter condition simply requires that any two elements of R should have a GCD in $R[\vartheta]$ which can be expressed as a linear combination of the elements.

THEOREM 2–18. *If $R(\vartheta)$ is of degree n, and A is an ideal of $R[\vartheta]$, then there exist elements $\alpha_1, \ldots, \alpha_n$ of $R[\vartheta]$ such that every element of*

A can be uniquely represented in the form

$$k_1\alpha_1 + \cdots + k_n\alpha_n, \qquad k_1, \ldots, k_n \text{ in } Z.$$

Remark: Note that the k's are rational integers, and not elements of $R[\vartheta]$. The numbers $\alpha_1, \ldots, \alpha_n$ of the theorem are called a *basis* of A; they may be taken as a set of generators of A, but may not be the smallest such set.

Proof: If the polynomial defining an element $\alpha \neq 0$ of A is $p(x)$, then for some h, the zeros of $p^h(x)$ are the field conjugates of α, so that

$$p^h(x) = x^n + a_1 x^{n-1} + \cdots \pm \mathbf{N}\alpha,$$

and $\mathbf{N}\alpha = \pm(\alpha^{n-1} + a_1\alpha^{n-2} + \cdots)\alpha$ is in A. Hence A contains a rational integer different from zero, and therefore a smallest positive integer, say a. If $\rho_1, \ldots, \rho_n$ is an integral basis of $R[\vartheta]$, then A contains $a\rho_i$ for each i. Let a_{11} be the smallest positive rational integer such that the number

$$\alpha_1 = a_{11}\rho_1$$

is in A. Since A contains $a_{11}\rho_1$ and $a\rho_2$, it contains numbers which are linear combinations of ρ_1 and ρ_2 with coefficients in Z. Of these there is one (not necessarily unique) for which the coefficient of ρ_2 is positive and minimal. Let it be

$$\alpha_2 = a_{21}\rho_1 + a_{22}\rho_2.$$

Similarly, for $\nu = 3, \ldots, n$, put

$$\alpha_\nu = a_{\nu 1}\rho_1 + a_{\nu 2}\rho_2 + \cdots + a_{\nu\nu}\rho_\nu,$$

where $a_{\nu i}$ is in Z for $1 \leq i \leq \nu$ and $a_{\nu\nu}$ is positive and minimal for α_ν in A. It is asserted that $\alpha_1, \ldots, \alpha_n$ form a basis of A.
Suppose that

$$\alpha = c_1\rho_1 + \cdots + c_n\rho_n, \qquad c_1, \ldots, c_n \text{ in } Z,$$

is in A. Then so also is $\alpha - c\alpha_n$ for every c in Z. Since

$$0 \leq c_n - a_{nn}\left[\frac{c_n}{a_{nn}}\right] < a_{nn},$$

it follows from the minimality of a_{nn} that in the representation of the number $\alpha - [c_n/a_{nn}]\alpha_n$, the coefficient of ρ_n is 0, so that

$$\alpha - \left[\frac{c_n}{a_{nn}}\right]\alpha_n = d_1\rho_1 + \cdots + d_{n-1}\rho_{n-1}, \qquad d_1, \ldots, d_n \text{ in } Z.$$

Repeating the argument, we find that

$$\alpha - \left[\frac{c_n}{a_{nn}}\right]\alpha_n - \left[\frac{d_{n-1}}{a_{n-1,n-1}}\right]\alpha_{n-1} = e_1\rho_1 + \cdots + e_{n-2}\rho_{n-2},$$

$$e_1, \ldots, e_{n-2} \text{ in } Z.$$

After n steps, we have

$$\alpha = \left[\frac{c_n}{a_{nn}}\right]\alpha_n + \left[\frac{d_{n-1}}{a_{n-1,n-1}}\right]\alpha_{n-1} + \cdots + \left[\frac{g_1}{a_{11}}\right]\alpha_1,$$

the desired representation.

If there were two representations of the same number, their difference would be a nontrivial representation of 0:

$$k_1\alpha_1 + \cdots + k_n\alpha_n = 0, \qquad k_1^2 + \cdots + k_n^2 > 0.$$

But then also

$$k_1\alpha_1^{(m)} + \cdots + k_n\alpha_n^{(m)} = 0, \qquad m = 1, 2, \ldots, n,$$

which implies that $\Delta(\alpha_1, \ldots, \alpha_n) = 0$, contrary to the equation

$$\Delta(\alpha_1, \ldots, \alpha_n) = a_{11}^2 a_{22}^2 \cdots a_{nn}^2 \Delta(\rho_1, \ldots, \rho_n) \neq 0.$$

The proof is complete.

From their definitions, it is clear that each coefficient a_{ii} is positive and not larger than a, the smallest positive integer in A. We would like to show that bounds can also be put on the other coefficients a_{ij}, $1 \leq j < i \leq n$. We have

$$\alpha_1 = a_{11}\rho_1,$$

$$\alpha_2 = a_{21}\rho_1 + a_{22}\rho_2,$$

$$\alpha_3 = a_{31}\rho_1 + a_{32}\rho_2 + a_{33}\rho_3, \qquad (13)$$

$$\vdots$$

$$\alpha_n = a_{n1}\rho_1 + a_{n2}\rho_2 + a_{n3}\rho_3 + \cdots + a_{nn}\rho_n.$$

THEOREM 2–19. *Every ideal in $R[\vartheta]$ has a basis $\alpha_1, \ldots, \alpha_n$, given by (13), in which the numbers a_{ij} are rational integers with*

$$0 \leq a_{ij} < a_{jj} \leq a_{11}.$$

Proof: It is clear that any system of numbers $\alpha_1, \ldots, \alpha_{i-1}$, $\alpha_i - k\alpha_j, \alpha_{i+1}, \ldots, \alpha_n$, in which k is a rational integer and $j \neq i$, is

also a basis of A. For if α is in A and

$$\alpha = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_n\alpha_n, \qquad k_1, \ldots, k_n \text{ in } Z,$$

then

$$\alpha = k_1\alpha_1 + \cdots + (k_j + kk_i)\alpha_j + \cdots + k_i(\alpha_i - k\alpha_j) + \cdots + k_n\alpha_n.$$

In the set of equations (13), subtract a suitable multiple of α_{n-1} from α_n, so that the new coefficient of ρ_{n-1} is non-negative but smaller than $a_{n-1,n-1}$. Then subtract a suitable multiple of α_{n-2}, so that the new coefficient of ρ_{n-2} is smaller than $a_{n-2,n-2}$; this does not disturb the coefficient of ρ_{n-1}. Continuing the process, we come eventually to a basis element $\alpha_n{}'$ such that $0 \leq a_{ni}{}' < a_{ii}$, for $i = 1, \ldots, n - 1$. Then we change α_{n-1} by subtracting off suitable multiples of α_{n-2}, $\alpha_{n-3}, \ldots, \alpha_1$, etc. The result is a basis as described in the theorem.

COROLLARY. *A positive rational integer occurs in only finitely many ideals of $R[\vartheta]$.*

This follows immediately from the theorem, for if a is in A, then $a_{11} \leq a$, and there are only finitely many sets of coefficients a_{ij} satisfying the conditions of the theorem.

The discriminant of the elements of a basis of an ideal is called the *discriminant of the ideal;* its value is independent of the choice of basis. For if $\alpha_1, \ldots, \alpha_n$ and $\alpha_1{}', \ldots, \alpha_n{}'$ are bases of A, then there are h_{kl} in Z such that

$$\alpha_k = \sum_{l=1}^{n} h_{kl}\alpha_l{}', \qquad k = 1, \ldots, n,$$

and

$$\det |h_{kl}| \neq 0.$$

Hence

$$\Delta(\alpha_1, \ldots, \alpha_n) = (\det |h_{kl}|)^2 \Delta(\alpha_1{}', \ldots, \alpha_n{}'),$$

so that the discriminants have the same sign and

$$\Delta(\alpha_1{}', \ldots, \alpha_n{}') | \Delta(\alpha_1, \ldots, \alpha_n).$$

By symmetry,

$$\Delta(\alpha_1, \ldots, \alpha_n) | \Delta(\alpha_1{}', \ldots, \alpha_n{}'),$$

and the discriminants are equal.

1. Show that every ideal in Z is principal.

2. If $A = [p, a + b\sqrt{d}]$ is an ideal of $R[\sqrt{d}]$, where p is a rational prime and d is a square-free integer not of the form $4k + 1$, show that p and $a + (b - p[b/p])\sqrt{d}$ form a basis for A.

2–6 The arithmetic of ideals. Ideals are special kinds of sets of elements. The emphasis so far has been on the elements comprising the sets. The whole power of the theory of ideals, however, lies in considering them not as collections of elements, but as entities in their own right, which can be combined according to certain operations.

The first of these operations is multiplication. If $A = [\alpha_1, \ldots, \alpha_r]$ and $B = [\beta_1, \ldots, \beta_s]$, then the *product* AB is the ideal

$$[\alpha_1\beta_1, \ldots, \alpha_1\beta_s, \alpha_2\beta_1, \ldots, \alpha_r\beta_s].$$

The product ideal does not depend on the representation chosen for A and B. To show this, let $AB = C$, and suppose that also

$$A = [\alpha_1{'}, \ldots, \alpha_{r'}'], \qquad B = [\beta_1{'}, \ldots, \beta_{s'}'].$$

To keep matters straight, designate these last ideals by A' and B', even though they are equal to A and B. We must show that every element of C is also an element of $A'B' = C'$, and conversely.

First of all, $\alpha_i{'}$ is in A and $\beta_j{'}$ is in B, so that we can write

$$\alpha_i{'} = \lambda_1\alpha_1 + \cdots + \lambda_r\alpha_r, \qquad \beta_j{'} = \mu_1\beta_1 + \cdots + \mu_s\beta_s.$$

Hence the number

$$\alpha_i{'}\beta_j{'} = \sum \lambda_k \mu_l \alpha_k \beta_l = \sum \nu_{kl} \alpha_k \beta_l$$

is in C for $1 \leq i \leq r'$, $1 \leq j \leq s'$. Since C is an ideal, every linear combination of the numbers $\alpha_i{'}\beta_j{'}$ is in C; thus C' is a subset of C. Hence $C = C'$, by symmetry.

THEOREM 2–20. *If A is an ideal of $R[\vartheta]$, there exists an ideal B of $R[\vartheta]$ such that AB is a principal ideal $[a]$, where a is in Z.*

Remark: It is this theorem which is the crux of the whole matter. As indicated in the discussion at the beginning of Section 2–5, we are trying to enlarge the set of possible divisors of an integer by introducing ideal elements. Given any such divisor, there should certainly be a second divisor whose product with the first is the original integer.

Since we have taken divisors as sets, we must identify the original integer with the set of all its multiples. It should be noted that all the associates of a given integer generate the same principal ideal.

Proof: Suppose $A = [\alpha_1, \ldots, \alpha_r]$, and put

$$f(x) = \alpha_1 + \alpha_2 x + \cdots + \alpha_r x^{r-1}.$$

By representing $\alpha_1, \ldots, \alpha_r$ as polynomials in ϑ, and replacing ϑ in all the polynomials by $\vartheta_2, \vartheta_3, \ldots, \vartheta_n$ in turn, we get sets $\alpha_1^{(\nu)}, \ldots, \alpha_r^{(\nu)}$, where $\nu = 2, 3, \ldots, n$. We define

$$g(x) = \prod_{\nu=2}^{n} (\alpha_1^{(\nu)} + \alpha_2^{(\nu)} x + \cdots + \alpha_r^{(\nu)} x^{r-1})$$

$$= \beta_1 + \beta_2 x + \cdots + \beta_s x^{s-1}.$$

The β's are symmetric polynomials, with rational integral coefficients, in all the conjugates of $\alpha_1, \ldots, \alpha_r$ except $\alpha_1, \ldots, \alpha_r$ themselves. Hence they are polynomials in $\alpha_1, \ldots, \alpha_r$, with coefficients in Z, and therefore are in $R[\vartheta]$. It is asserted that the ideal $B = [\beta_1, \ldots, \beta_s]$ satisfies the conditions of the theorem.

Put

$$f(x)g(x) = \gamma_1 + \gamma_2 x + \cdots + \gamma_{r+s-1} x^{r+s-2}.$$

Since each γ is a symmetric polynomial, with rational integral coefficients, in each α_i and its conjugates, the γ's are themselves rational integers. Let their GCD be a. Then a can be represented as a linear combination of $\gamma_1, \ldots, \gamma_{r+s-1}$, with coefficients in Z; since $\gamma_1, \ldots, \gamma_{r+s-1}$ are obviously in AB, a is in AB, and so $[a]$ is a subset of AB.

If we knew that a divides every product $\alpha_i \beta_j$, then we would know that every element of AB is contained in $[a]$. The proof will therefore be complete when we prove Theorem 2–21, which is A. Hurwitz' extension of a theorem due to Gauss.

THEOREM 2–21. *Let*

$$A(x) = \alpha_0 x^r + \cdots + \alpha_r, \qquad B(x) = \beta_0 x^s + \cdots + \beta_s,$$

where $\alpha_0 \beta_0 \neq 0$, be polynomials with integral algebraic coefficients. If an algebraic integer δ divides every coefficient of

$$C(x) = A(x)B(x) = c_0 x^t + \cdots + c_t,$$

in the sense that each quotient c_i/δ is an algebraic integer, then δ also divides every product $\alpha_k \cdot \beta_l$.

Proof: First we prove a lemma: *if*

$$f(x) = \delta_0 x^u + \cdots + \delta_u, \qquad \delta_0 \neq 0,$$

is any polynomial with integral algebraic coefficients and a zero ρ, then $f(x)/(x - \rho)$ has integral coefficients. The proof is by induction on u. If $u = 1$, then $f(x) = \delta_0 x + \delta_1$, and

$$\frac{f(x)}{x - \rho} = \frac{\delta_0 x + \delta_1}{x + \delta_1/\delta_0} = \delta_0$$

is an integer. Suppose the lemma true for all polynomials of degree less than u. Then the polynomial

$$Q(x) = f(x) - \delta_0 x^{u-1}(x - \rho)$$

has integral algebraic coefficients (by the second part of Theorem 2–10), and has degree less than u and vanishes for $x = \rho$. By the induction hypothesis,

$$\frac{Q(x)}{x - \rho} = \frac{f(x)}{x - \rho} - \delta_0 x^{u-1}$$

has integral algebraic coefficients, and the same is therefore true of $f(x)/(x - \rho)$. The lemma follows by the induction principle.

By repeated application of the lemma, we deduce that if $f(x) = \delta_0(x - \rho_1) \cdots (x - \rho_u)$, then any product $\delta_0 \rho_1 \cdots \rho_k$ is an integer.

Returning to Theorem 2–21, suppose that

$$A(x) = \alpha_0(x - \rho_1) \cdots (x - \rho_r),$$
$$B(x) = \beta_0(x - \sigma_1) \cdots (x - \sigma_s).$$

By assumption, the polynomial

$$\frac{C(x)}{\delta} = \frac{\alpha_0 \beta_0}{\delta}(x - \rho_1) \cdots (x - \sigma_s)$$

has integral coefficients, and it follows that any product

$$\frac{\alpha_0 \beta_0}{\delta} \rho_{n_1} \cdots \rho_{n_i} \sigma_{m_1} \cdots \sigma_{m_j}, \qquad \begin{cases} 1 \leq n_1 < n_2 < \cdots < n_i \leq r, \\ 1 \leq m_1 < m_2 < \cdots < m_j \leq s, \end{cases} \qquad (14)$$

is an integer. Since α_k/α_0 and β_l/β_0 are elementary symmetric func-

tions in the ρ's and σ's, respectively, the number

$$\frac{\alpha_k \beta_l}{\delta} = \frac{\alpha_0 \beta_0}{\delta} \cdot \frac{\alpha_k}{\alpha_0} \cdot \frac{\beta_l}{\beta_0}$$

is a sum of terms of the form (14), and is therefore an integer. The proof is complete.

THEOREM 2-22. *If $AC = BC$, then $A = B$.*

Remark: Note that there is no zero ideal.

Proof: Let D be an ideal such that $CD = [e]$, a principal ideal. Then $ACD = BCD$, so $A[e] = B[e]$. Thus e times any element of A is equal to e times some element of B, and so $A = B$.

If $A = BC$, then we say that C *divides* A, and write $C|A$.

THEOREM 2-23. *$A|C$ if and only if every element of C is in A.*

Proof: If $A = [\alpha_1, \ldots, \alpha_r]$ and $B = [\beta_1, \ldots, \beta_s]$, then $AB = C = [\ldots, \alpha_i\beta_j, \ldots]$, so every element of C is in A, and also in B.

Conversely, suppose that every element of C is in A. Then every element of CD is in AD, for every D. Choose D so that $AD = [e]$ is principal, and let $CD = [\sigma_1, \ldots, \sigma_t]$. Then for each i with $1 \leq i \leq t$, $\sigma_i = e\lambda_i$ for a suitable integer λ_i. Hence $CD = [e][\lambda_1, \ldots, \lambda_t] = AD[\lambda_1, \ldots, \lambda_t]$, and by Theorem 2-22, $C = A[\lambda_1, \ldots, \lambda_t]$, so that $A|C$.

THEOREM 2-24. *An ideal is divisible by only a finite number of ideals.*

Proof: If the ideal is A, choose B so that $AB = [c]$, where c is a positive integer. Then c is in A and in every divisor of A, and by the corollary to Theorem 2-19, there are only finitely many such ideals.

A common divisor of A and B which is divisible by every common divisor is called a *greatest common divisor* (GCD) of A and B.

THEOREM 2-25. *Every pair of ideals A and B has a unique GCD, (A, B). It is composed of the numbers $\alpha + \beta$, where α runs over A and β over B.*

Proof: Let D be the set described in the theorem; it is clearly an ideal. Since 0 is in A and B, D contains every element of A and of B, and so is a divisor of A and of B. Any common divisor of A and B

contains all the elements of A and of B, and since it is closed under addition, it contains all numbers $\alpha + \beta$, and so divides D.

If D' is also a GCD of A and B, then D and D' are divisors of each other, and so each contains the other. Thus $D = D'$.

If the GCD of A and B is [1], we say that A and B are *relatively prime*. As an immediate consequence of this definition and Theorem 2–25, we have

THEOREM 2–26. *If A and B are relatively prime, there exist α in A and β in B such that $\alpha + \beta = 1$.*

THEOREM 2–27. *If $A|BC$ and A is prime to B, it divides C.*

Proof: Choose α in A and β in B so that $\alpha + \beta = 1$. Then if γ is in C, $\alpha\gamma + \beta\gamma = \gamma$, and $\beta\gamma$ and $\alpha\gamma$ are in A, so that γ is in A. Hence $A|C$.

If A has no divisors except itself and [1], then A is said to be *prime*.

THEOREM 2–28. *Every ideal can be represented as a finite product of prime ideals, and the representation is unique except for the order of factors.*

The finiteness of the representation follows from Theorem 2–24, and the uniqueness from Theorem 2–27.

In particular, it follows that the principal ideal generated by any element of $R[\vartheta]$ has a unique factorization into prime ideals of $R[\vartheta]$. If these prime factors are themselves always principal ideals, we might expect that ideals can be dispensed with entirely, and that there is then unique factorization of the numbers themselves.

THEOREM 2–29. *A necessary and sufficient condition that $R[\vartheta]$ be a unique factorization domain is that every ideal of $R[\vartheta]$ be a principal ideal.*

Proof: Uniqueness of factorization in $R[\vartheta]$ is equivalent to the property:

$$\text{if } \alpha|\beta\gamma \text{ and } \alpha \text{ and } \beta \text{ are relatively prime, then } \alpha|\gamma. \tag{15}$$

For if the domain has this property, unique factorization can be proved in the usual way, while if factorization is unique and $\alpha|\beta\gamma$, then every prime π dividing α must occur in the factorization of $\beta\gamma$; since this

factorization is the product of the factorizations of β and γ, if π does not occur in β it must occur in γ.

Suppose that factorization is unique in $R[\vartheta]$, so that (15) holds. Then if π is a prime number, $[\pi]$ is a prime ideal. For if $[\pi] = AB$, where neither A nor B is $[\pi]$, there would exist an α in A and a β in B, neither of which is divisible by π, while their product is.

Let P be any prime ideal, and $\alpha = \pi_1{}^{n_1} \dots \pi_r{}^{n_r}$ any element of P. Then

$$[\alpha] = [\pi_1]^{n_1} \dots [\pi_r]^{n_r},$$

and since α is in P, so is every element of $[\alpha]$, whence $P|[\alpha]$ and P is one of the principal ideals $[\pi_k]$. Since every prime ideal is principal, every ideal is principal.

Now suppose that every ideal in $R[\vartheta]$ is principal, and that α and β are relatively prime. Then $[\alpha, \beta] = [\gamma]$, for some γ, and every linear combination $\lambda\alpha + \mu\beta$ is a multiple of γ. Taking $\lambda = 1$ and $\mu = 0$, we have $\gamma|\alpha$; for $\lambda = 0$ and $\mu = 1$, we obtain $\gamma|\beta$. Hence γ is a unit, $[\alpha, \beta] = [1]$, and we can take $\gamma = 1$. Thus there are λ and μ such that $\lambda\alpha + \mu\beta = 1$, so that if $\alpha|\beta\gamma$, then α divides $\lambda\alpha\gamma + \mu\beta\gamma = \gamma$, and (15) holds. Hence factorization is unique.

PROBLEMS

1. Using Theorem 2–21, reformulate and prove the new version of Eisenstein's irreducibility criterion, as given in Problem 1, Section 2–2.

2. Show that if $A = [a_1 + b_1\sqrt{d}, a_2 + b_2\sqrt{d}]$ is an ideal of $R[\sqrt{d}]$, then the product of A with its *conjugate ideal* $A' = [a_1 - b_1\sqrt{d}, a_2 - b_2\sqrt{d}]$ is principal.

2–7 Congruences. The norm of an ideal. Two elements α and β of $R[\vartheta]$ will be said to be *congruent modulo an ideal A* if their difference lies in A, that is, if A divides the ideal $[\alpha - \beta]$. This is a natural extension of the earlier notion of congruence of rational integers, if the modulus m is identified with the principal ideal $[m]$. The familiar properties of congruences are easily seen to hold.

For fixed α, the set of all elements of $R[\vartheta]$ which are congruent to α modulo A is called a *residue class modulo A*.

THEOREM 2–30. *There are only finitely many residue classes modulo A.*

Proof: Choose B so that $AB = [c]$, where c is a rational integer. Then $\alpha_1 \not\equiv \alpha_2 \pmod{A}$ implies that $\alpha_1 \not\equiv \alpha_2 \pmod{[c]}$, since $A \,|\, [c]$ and therefore A contains $[c]$. So if we can show that there are only finitely many elements, no two of which are congruent modulo $[c]$, the theorem follows. But this is an immediate consequence of the fact that in the basis representation

$$\alpha = r_1\omega_1 + \cdots + r_n\omega_n,$$

where $\omega_1, \ldots, \omega_n$ form an integral basis of $R[\vartheta]$, each of the rational integral coefficients $r_1, \ldots, r_n$ has only c possible values modulo c, and that if

$$r_i \equiv r_i' \pmod{c}, \qquad i = 1, \ldots, n,$$

then

$$r_1\omega_1 + \cdots + r_n\omega_n \equiv r_1'\omega_1 + \cdots + r_n'\omega_n \pmod{[c]}.$$

The number of residue classes modulo A is called the *norm* of A, written $\mathbf{N}A$. For the time being, it is necessary to distinguish between $\mathbf{N}\alpha$ and $\mathbf{N}[\alpha]$, the norms of the number α and the ideal $[\alpha]$, respectively. However, we shall soon see that the two quantities are essentially the same.

THEOREM 2–31. *If $R(\vartheta)$ has discriminant Δ, and A is an ideal of $R[\vartheta]$ having discriminant $\Delta(A)$, then*

$$\Delta(A) = (\mathbf{N}A)^2\Delta.$$

Proof: Let $\alpha_1, \cdots, \alpha_n$ be the basis of A described in Theorem 2–19, and let $\rho_1, \cdots, \rho_n$ be a basis of $R[\vartheta]$. Then

$$\Delta(A) = (a_{11} \cdots a_{nn})^2\Delta,$$

and we must show that $\mathbf{N}A = a_{11} \cdots a_{nn}$; that is, that there are $a_{11} \cdots a_{nn}$ numbers of $R[\vartheta]$, no two of which are congruent modulo A and such that every element of $R[\vartheta]$ is congruent to one of them. We show that this is true of the numbers

$$r_1\rho_1 + \cdots + r_n\rho_n,$$

where $0 \le r_k < a_{kk}$ for $k = 1, \ldots, n$. If two of these numbers are congruent, say

$$r_1\rho_1 + \cdots + r_n\rho_n \equiv r_1'\rho_1 + \cdots + r_n'\rho_n \pmod{A},$$

and $r_n \ge r_n'$, then

$$(r_1 - r_1')\rho_1 + \cdots + (r_n - r_n')\rho_n \equiv 0 \pmod{A}.$$

But a_{nn} is the smallest positive rational integer for which any number of the form

$$s_1\rho_1 + \cdots + s_{n-1}\rho_{n-1} + a_{nn}\rho_n$$

is in A; since $0 \leq r_n - r_n' < a_{nn}$, it follows that $r_n - r_n' = 0$. Similarly, $r_{n-1} = r_{n-1}', \ldots, r_1 = r_1'$.

If

$$\beta = s_1\rho_1 + \cdots + s_n\rho_n,$$

then

$$\beta - \left[\frac{s_n}{a_{nn}}\right]\alpha_n = s_1'\rho_1 + \cdots + s_{n-1}'\rho_{n-1} + b_n\rho_n,$$

where $0 \leq b_n < a_{nn}$. By iteration,

$$\beta - \left[\frac{s_n}{a_{nn}}\right]\alpha_n - \left[\frac{s_{n-1}'}{a_{n-1,n-1}}\right]\alpha_{n-1} - \cdots = b_1\rho_1 + \cdots + b_n\rho_n,$$

where $0 \leq b_k < a_{kk}$ for $k = 1, \ldots, n$, and

$$\beta \equiv b_1\rho_1 + \cdots + b_n\rho_n \pmod{A}.$$

COROLLARY. $\mathbf{N}[\alpha] = |\mathbf{N}\alpha|$.

For $\alpha\rho_1, \ldots, \alpha\rho_n$ is clearly a basis for $[\alpha]$, and

$$\Delta(\alpha\rho_1, \ldots, \alpha\rho_n) = (\mathbf{N}\alpha)^2\Delta,$$

so that $(\mathbf{N}\alpha)^2 = (\mathbf{N}[\alpha])^2$. But $\mathbf{N}[\alpha]$, being the number of residue classes, is positive.

THEOREM 2-32. *If A and B are ideals, then there is an α in A such that $([\alpha], AB) = A$.*

If such an α exists, then clearly $[\alpha] = AC$, where $(B, C) = [1]$. If we rephrase the theorem, its close relation to Theorem 2-20 becomes clear: given two ideals A and B, there is a C such that AC is principal and $(B, C) = [1]$.

Proof: Let $P_1, \ldots, P_r$ be the distinct primes dividing AB, and let

$$A = P_1^{e_1} \cdots P_r^{e_r}, \qquad e_i \geq 0.$$

Put

$$D_i = \prod_{\substack{1 \leq j \leq r \\ j \neq i}} P_j^{e_j+1}, \qquad i = 1, \ldots, r.$$

Since $(D_1, \ldots, D_r) = [1]$, there are numbers δ_i in D_i, for $i = 1, \ldots, r$, such that

$$\delta_1 + \cdots + \delta_r = 1.$$

Then $[\delta_i]$ is divisible by D_i, and therefore by P_k for $k \neq i$, and therefore not by P_i, since 1 is not. Now let α_i be an element of $P_i^{e_i}$ which does not occur in $P_i^{e_i+1}$, for $i = 1, \ldots, r$, and put

$$\alpha = \alpha_1 \delta_1 + \cdots + \alpha_r \delta_r.$$

Then for each i, every term but one in this representation of α occurs in $P_i^{e_i+1}$, while the remaining term occurs in $P_i^{e_i}$ but not in $P_i^{e_i+1}$. Hence $A | [\alpha]$, but

$$\left(\frac{[\alpha]}{A}, B \right) = [1].$$

THEOREM 2–33. *The congruence*

$$\alpha \xi \equiv \beta \pmod{A}$$

is solvable if and only if $D | [\beta]$, where $D = ([\alpha], A)$. The solution, if it exists, is unique modulo A/D.

Proof: If ξ is a solution, then $\alpha \xi - \beta = \gamma$ is in A, and therefore in D. Since also $\alpha \xi$ is in D, it follows that β is in D, so $D | [\beta]$.

If β is in D, then it is the sum of an element of $[\alpha]$ and an element of A; that is, $\beta = \alpha \xi + \delta$. Since $\delta \equiv 0 \pmod{A}$, $\beta \equiv \alpha \xi \pmod{A}$.

If $\alpha \xi \equiv \alpha \xi' \equiv \beta \pmod{A}$, then $\alpha(\xi - \xi') \equiv 0 \pmod{A}$. Hence if $[\alpha] = DA_1$ and $A = DA_2$, then $(A_1, A_2) = [1]$ and

$$DA_2 | DA_1 [\xi - \xi'],$$

$$A_2 | A_1 [\xi - \xi'],$$

$$\xi \equiv \xi' \pmod{A_2}.$$

THEOREM 2–34. $\mathbf{N}(AB) = \mathbf{N}A \cdot \mathbf{N}B.$

Proof: By Theorem 2–32, there is a γ such that

$$([\gamma], AB) = A.$$

Let $\mathbf{N}A = n_1$, $\mathbf{N}B = n_2$, and let $\alpha_1, \ldots, \alpha_{n_1}$ and $\beta_1, \ldots, \beta_{n_2}$ be complete residue systems modulo A and B, respectively. We shall show that the $n_1 n_2$ numbers $\alpha_i + \gamma \beta_j$ form a complete residue system modulo AB.

If
$$\alpha_i + \gamma\beta_j \equiv \alpha_k + \gamma\beta_l \pmod{AB},$$

then
$$\gamma(\beta_j - \beta_l) \equiv \alpha_k - \alpha_i \pmod{AB},$$

and by Theorem 2-33, $([\gamma], AB)|[\alpha_k - \alpha_i]$, so that $A|[\alpha_k - \alpha_i]$. But this gives $\alpha_k \equiv \alpha_i \pmod{A}$, so $k = i$. Moreover,

$$\gamma(\beta_j - \beta_l) \equiv 0 \pmod{AB},$$

$$\beta_j - \beta_l \equiv 0 \pmod{B},$$

$$j = l.$$

To show that every integer δ is congruent to one of the above numbers, choose α_i so that $\delta \equiv \alpha_i \pmod{A}$. Then the congruence

$$\gamma\xi \equiv \delta - \alpha_i \pmod{AB}$$

is solvable, since $([\gamma], AB) = A$ is a divisor of $[\delta - \alpha_i]$. Finally, ξ is unique modulo B, and can therefore be taken to be one of the numbers β_j.

THEOREM 2-35. $\mathbf{N}A$ *is an element of* A.

Proof: If $\alpha_1, \ldots, \alpha_{\mathbf{N}A}$ is a complete residue system modulo A, then so is $\alpha_1 + 1, \ldots, \alpha_{\mathbf{N}A} + 1$. Hence

$$\alpha_1 + \cdots + \alpha_{\mathbf{N}A} \equiv (\alpha_1 + 1) + \cdots + (\alpha_{\mathbf{N}A} + 1) \pmod{A},$$

$$0 \equiv \mathbf{N}A \pmod{A}.$$

COROLLARY. *There are only finitely many ideals of given norm.*

For by the corollary to Theorem 2-19, a positive rational integer occurs in only finitely many ideals.

PROBLEMS

1. Show that if P is a prime ideal of $R[\vartheta]$, the congruence

$$x^m + \alpha_1 x^{m-1} + \cdots + \alpha_m \equiv 0 \pmod{P}$$

with coefficients in $R[\vartheta]$ has at most m incongruent solutions modulo P.

2. Show that if P is a prime ideal of $R[\vartheta]$, α is an element of $R[\vartheta]$, and $P \nmid [\alpha]$, then

$$\alpha^{\mathbf{N}P-1} \equiv 1 \pmod{P}.$$

2–8 Prime ideals

THEOREM 2–36.   *If* $\mathbf{N}A$ *is prime, so is* A.

This follows immediately from Theorem 2–34.

THEOREM 2–37.   *There are infinitely many prime ideals P in any domain $R[\vartheta]$. Each such P divides exactly one rational prime p, and $\mathbf{N}P = p^f$, where f, called the* degree *of P, is a positive integer not exceeding the degree of $R(\vartheta)$.*

Proof: Let p be a rational prime, and let P be one of the factors of $[p]$ in $R[\vartheta]$. Then if P also divided the ideal defined by another rational prime p', it would divide their GCD, which is $[1]$. Hence each P divides at most one p, and each of the infinitely many rational primes p is divisible by at least one P, so that there must be infinitely many P's.

Now let a be a rational integer such that $P|[a]$; by Theorem 2–35, we could take $a = \mathbf{N}P$. If $a = p_1 \cdots p_r$, then

$$P|[p_1] \cdots [p_r],$$

and so $P|[p_i]$ for some i.

Finally, if $P|[p]$ then $[p] = PA$ for some A. By the corollary to Theorem 2–31,

$$\mathbf{N}[p] = |\mathbf{N}p| = p^n,$$

and so $\mathbf{N}(PA) = \mathbf{N}P \cdot \mathbf{N}A = p^n$. Hence $\mathbf{N}P|p^n$, and the proof is complete.

Theorem 2–37 shows that the primes of $R[\vartheta]$ are to be found among the factors of the principal ideals $[p]$. Only partial information is available about the way these ideals decompose, and the derivation of most of what is known is too intricate for inclusion here, but we can prove the simpler half of a famous theorem due to Dedekind, which states that $[p]$ *is divisible by the square of a prime ideal in $R[\vartheta]$ if and only if p divides Δ, the discriminant of $R(\vartheta)$.*

THEOREM 2–38.   *If p does not divide Δ, then $[p]$ factors as a product of (one or more) distinct prime ideals.*

Proof: Suppose that $P^2|[p]$, so that $[p] = P^2M$. Choose an element α of PM which does not belong to P^2M, so that $p|\alpha^2$ but $p\nmid\alpha$. Since $p \geq 2$, $p|(\alpha\beta)^p$ for every β in $R[\vartheta]$.

For an arbitrary element γ of $R(\vartheta)$, define $\mathsf{S}\gamma$, the *trace* of γ, by the equation

$$\mathsf{S}\gamma = \gamma' + \cdots + \gamma^{(n)},$$

where $\gamma', \ldots, \gamma^{(n)}$ are the field conjugates of γ. By the Symmetric Function Theorem, $\mathsf{S}\gamma$ is in Z if γ is an integer, and it is clear that $\mathsf{S}(r\gamma) = r\mathsf{S}\gamma$ if r is rational. In particular,

$$\mathsf{S}\frac{(\alpha\beta)^p}{p} = \frac{\mathsf{S}(\alpha\beta)^p}{p}$$

is in Z, so that $\mathsf{S}(\alpha\beta)^p$ is in $[p]$. By the multinomial theorem, if $\beta', \ldots, \beta^{(n)}$ are the field conjugates of β, then

$$(\mathsf{S}(\alpha\beta))^p = (\alpha'\beta' + \cdots + \alpha^{(n)}\beta^{(n)})^p$$
$$\equiv (\alpha'\beta')^p + \cdots + (\alpha^{(n)}\beta^{(n)})^p = \mathsf{S}((\alpha\beta)^p)$$
$$\equiv 0 \pmod{p},$$

and since $\mathsf{S}(\alpha\beta)$ is a rational integer, $p|\mathsf{S}(\alpha\beta)$.

Now let $\rho_1, \ldots, \rho_n$ be an integral basis for $R[\vartheta]$. Then

$$\alpha = h_1\rho_1 + \cdots + h_n\rho_n,$$

where the h's are rational integers not all divisible by p, since $p\nmid\alpha$. For $1 \leq i \leq n$ we have

$$\mathsf{S}(\alpha\rho_i) = \mathsf{S}\left(\sum_{j=1}^{n} h_j\rho_j\rho_i\right) = \sum_{j=1}^{n} h_j\mathsf{S}(\rho_i\rho_j).$$

Let

$$d = \det |\mathsf{S}(\rho_i\rho_j)|,$$

and let A_{ij} be the cofactor of $\mathsf{S}(\rho_i\rho_j)$ in d. Then for $k = 1, 2, \ldots, n$,

$$\sum_{i=1}^{n} A_{ik} \sum_{j=1}^{n} h_j\mathsf{S}(\rho_i\rho_j) = \sum_{j=1}^{n} h_j \sum_{i=1}^{n} A_{ik}\mathsf{S}(\rho_i\rho_j) = dh_k.$$

Since

$$p\bigg|\sum_j h_j\mathsf{S}(\rho_i\rho_j)$$

for each i, it is also true that $p|dh_k$ for each k; p therefore divides d. Finally,

$$d = \det |\mathsf{S}(\rho_i\rho_j)| = \det |\sum_k \rho_i^{(k)}\rho_j^{(k)}| = \det |\rho_i^{(k)}|^2 = \Delta;$$

hence $p|\Delta$.

As an illustration of the present theorem, note that in the field $R(i)$, of discriminant -4, we have

$$[2] = [1 + i]^2,$$
$$[p] = [a + bi][a - bi], \quad \text{if } a^2 + b^2 = p \equiv 1 \pmod 4,$$
$$[q] = P, \text{ a prime ideal}, \quad \text{if } q \equiv 3 \pmod 4.$$

Here $[1 + i]$, $[a + bi]$, and $[a - bi]$ are prime ideals of degree 1, while each P is of degree 2.*

THEOREM 2–39. *Each ideal $[p]$, where p is a rational prime, splits into at most n ideal factors in the integral domain of a field of degree n.*

Proof: If $\quad [p] = P_1 \cdots P_r,$

then $\quad p^n = \mathbf{N}[p] = \mathbf{N}P_1 \cdots \mathbf{N}P_r,$

and for each i, $\mathbf{N}P_i > 1$. Hence $r \leq n$.

PROBLEMS

1. In the domain $R[\sqrt{-5}]$, put

$$A = [3, 4 + \sqrt{-5}], \quad B = [3, 4 - \sqrt{-5}], \quad C = [7, 4 + \sqrt{-5}],$$
$$D = [7, 4 - \sqrt{-5}].$$

Show that $AB = [3]$, $CD = [7]$, $AC = [4 + \sqrt{-5}]$, $BD = [4 - \sqrt{-5}]$, and that A, B, C, and D are prime ideals. Factor $[1 + 2\sqrt{-5}]$.

2. Let $R(\sqrt{d})$, where d is square-free, have discriminant Δ. If q is an odd prime dividing Δ, show that the ideals

$$\left[q, \frac{\Delta + \sqrt{\Delta}}{2}\right] \quad \text{and} \quad \left[q, \frac{\Delta - \sqrt{\Delta}}{2}\right]$$

are equal, and that their product is q. Show also that if Δ is even, then $[2] = [2, \sqrt{d}]^2$ for $d \equiv 2 \pmod 4$ and that $[2] = [2, 1 + \sqrt{d}]^2$ if $d \equiv 3 \pmod 4$. This completes the proof of Dedekind's theorem, stated just before Theorem 2–38, in the case of a quadratic field.

2–9. Units of algebraic number fields. We saw in Section 2–4 that the units of a quadratic field $R(\sqrt{d})$ are determined by the solutions of the Pell equation with $N = \pm 1$ or ± 4, and it is an easy conse-

*Compare the remark following Theorem 7–7, Volume I.

quence of this relationship and standard properties of Pell's equation*
that the group of units of $R(\sqrt{d})$ has a *basis*, consisting of -1 and
the fundamental solution ϵ of the appropriate Pell equation. That is,
every unit of $R(\sqrt{d})$ can be written in the form $(-1)^\alpha \epsilon^\beta$, where α is
0 or 1 and β ranges over Z. We shall now show that this property is
not peculiar to quadratic fields, but that in fact the group of units in
each algebraic number field has a finite basis. (In general, if G is a
commutative multiplicative group and $b_1, \ldots, b_m$ are elements of G,
they are said to form a basis for G if every element of G can be repre-
sented in the form $b_1^{n_1} \cdots b_m^{n_m}$, and in every such representation of
the unit element e of G, the factor $b_i^{n_i} = e$ for $1 \le i \le m$.) This
theorem, which is due to Dirichlet, can be sharpened by giving
the exact number of basis elements, but for many purposes, including
the application to be made in the next chapter, the finiteness of the
basis suffices. The upper bound which we shall obtain is actually the
correct number.

We introduce the symbol $\lceil \alpha \rceil$ to designate the maximum of the
absolute values of the conjugates of the algebraic number α, and
denote by K a fixed algebraic number field.

THEOREM 2-40. *If a is a fixed positive number, there are only
finitely many integers α of K such that*

$$\lceil \alpha \rceil \le a.$$

If all conjugates of α have absolute value 1, then α is a root of unity.

Proof: If $\lceil \alpha \rceil \le a$ and deg $\alpha = n$, then each of the elementary
symmetric functions in α and its conjugates is numerically smaller
than some bound depending only on a and n. If α is an integer in K,
then n cannot exceed the degree of K, so that there are available only
finitely many coefficients for the defining polynomial of α, and there
are, therefore, only finitely many such α's in K.

If $|\alpha^{(i)}| = 1$ for $i = 1, \ldots, n$, then $\lceil \alpha^m \rceil = 1$ for all m in Z, so that
by what we have just proved, $\alpha^{m_1} = \alpha^{m_2}$ for some distinct exponents
m_1 and m_2. Hence $\alpha^{m_1 - m_2} = 1$, so that α is a root of unity.

THEOREM 2-41. *The group U of roots of unity in K is a finite cyclic
group.*

---

*See, for example, Theorems 8-5, 8-6, and 8-7, Volume I.

Proof: If ζ is a root of unity, then $|\zeta| = 1$, and the finiteness of the group follows from the preceding theorem. Let the various elements u_i of U be primitive w_i-th roots of unity, for $i = 1, \ldots, t$, and put

$$w = \max (w_1, \ldots, w_t).$$

For fixed i, the numbers $e^{2\pi i a/w_i}$ and $e^{2\pi i b/w}$ are in U, for every a and b in Z. If $(w_i, w) = d$, choose a and b so that $aw + bw_i = d$; then the product

$$e^{2\pi i(a/w_i + b/w)} = e^{2\pi i d/w_i w} = e^{2\pi i/\langle w_i, w \rangle}$$

is in U. It follows that the LCM of w_i and w does not exceed w, so that $w_i | w$ for $i = 1, \ldots, t$. Since the powers of

$$\zeta_0 = e^{2\pi i/w}$$

include all dth roots of unity if $d|w$, it is clear that ζ_0 generates U.

Now let ϑ, of degree n, be a primitive element of K, so that $K = R(\vartheta)$, and arrange the conjugates of ϑ in such an order that $\vartheta^{(1)}, \ldots, \vartheta^{(r_1)}$ are real, while $\vartheta^{(r_1+1)}, \ldots, \vartheta^{(n)}$ are not real. (Note that it is not necessarily true that $\vartheta^{(1)} = \vartheta$.) Then $n - r_1$ is an even number, say $2r_2$, and we can further order the nonreal conjugates so that $\vartheta^{(r_1+j)}$ and $\vartheta^{(r_1+r_2+j)}$ are complex-conjugate, for $j = 1, \ldots, r_2$. If α is any number of K, the field conjugates of α are such that $\alpha^{(1)}, \ldots, \alpha^{(r_1)}$ are real, while $\alpha^{(r_1+j)}$ and $\alpha^{(r_1+r_2+j)}$ are complex-conjugate for $j = 1, \ldots, r_2$. Of course some of these latter numbers may also be real, but in any case

$$|\alpha^{(r_1+j)}| = |\alpha^{(r_1+r_2+j)}|, \qquad \text{for } j = 1, \ldots, r_2. \tag{16}$$

If $\epsilon_1, \ldots, \epsilon_k$ are units of K, they are said to be *independent* if the relation

$$\epsilon_1^{a_1} \cdots \epsilon_k^{a_k} = 1, \qquad a_1, \ldots, a_k \text{ in } Z, \tag{17}$$

holds only for $a_1 = \cdots = a_k = 0$.

THEOREM 2–42. *Units $\epsilon_1, \ldots, \epsilon_k$ in K are independent if and only if the sole solution of the system*

$$\sum_{m=1}^{k} x_m \log |\epsilon_m^{(i)}| = 0, \qquad i = 1, 2, \ldots, r, \tag{18}$$

in rational integers is $x_1 = \cdots = x_k = 0$. Here $r = r_1 + r_2 - 1$.

Proof: Suppose that (17) has a solution in which not all the a's are zero. Then the analogous equation with each ϵ_m replaced by $\epsilon_m{}^{(i)}$ also holds, so that

$$\left|\epsilon_1{}^{(i)}\right|^{a_1} \cdots \left|\epsilon_k{}^{(i)}\right|^{a_k} = 1, \qquad i = 1, \ldots, n,$$

and

$$\sum_{m=1}^{k} a_m \log \left|\epsilon_m{}^{(i)}\right| = 0, \qquad i = 1, \ldots, n. \tag{19}$$

Conversely, if (19) holds with not all the rational integers $a_1, \ldots, a_k$ equal to zero, then $\epsilon_1{}^{a_1} \cdots \epsilon_k{}^{a_k}$ is an integer of K all of whose conjugates have absolute value 1; it is therefore a root of unity whose wth power is 1, and (17) holds with $a_1, \ldots, a_k$ replaced by $wa_1, \ldots, wa_k$. Hence the nontrivial solvability in Z of (19) is equivalent to the dependence of $\epsilon_1, \ldots, \epsilon_k$.

The truth of the theorem will now follow if we can prove that if the equations (19) hold with $i = 1, \ldots, r_1 + r_2 - 1$, then the remaining $n - r_1 - r_2 + 1$ equations are also correct. To show this, suppose that the first $r_1 + r_2 - 1$ equations are true, and define

$$e_i = \begin{cases} 1 & \text{for } 1 \leq i \leq r_1, \\ 2 & \text{for } r_1 + 1 \leq i \leq n. \end{cases}$$

Since each ϵ_m is a unit, its norm $\epsilon_m{}^{(1)} \cdots \epsilon_m{}^{(n)}$ has absolute value 1; by (16),

$$\sum_{i=1}^{n} \log \left|\epsilon_m{}^{(i)}\right| = \sum_{i=1}^{r_1+r_2} e_i \log \left|\epsilon_m{}^{(i)}\right| = 0.$$

Hence

$$\sum_{m=1}^{k} a_m \sum_{i=1}^{r_1+r_2} e_i \log \left|\epsilon_m{}^{(i)}\right| = \sum_{i=1}^{r_1+r_2} e_i \sum_{m=1}^{k} a_m \log \left|\epsilon_m{}^{(i)}\right| = 0,$$

so that

$$e_{r_1+r_2} \cdot \sum_{m=1}^{k} a_m \log \left|\epsilon_m{}^{(r_1+r_2)}\right| = - \sum_{i=1}^{r_1+r_2-1} e_i \sum_{m=1}^{k} a_m \log \left|\epsilon_m{}^{(i)}\right| = 0.$$

Thus (19) also holds for $i = r_1 + r_2$, and so, by (16), for

$$i = 1, 2, \ldots, n.$$

THEOREM 2–43. *If the relation* (18) *holds for some set of real numbers* $x_1, \ldots, x_k$ *which are not all zero, it also holds for rational integers* $\ldots, x_k$ *which are not all zero.*

Proof: Suppose the hypothesis fulfilled. Since the system (18) is certainly nontrivially solvable in rational integers if some ϵ_m is in U, it suffices to consider the case that all the units are of infinite order. Then each unit separately is independent. Now suppose that the units $\epsilon_1, \ldots, \epsilon_q$ are such that the equations

$$\sum_{m=1}^{q-1} \alpha_m \log |\epsilon_m^{(i)}| = 0, \qquad i = 1, \ldots, r, \qquad (20)$$

have the single real solution $\alpha_1 = \cdots = \alpha_{q-1} = 0$, while the system

$$\sum_{m=1}^{q} \alpha_m \log |\epsilon_m^{(i)}| = 0, \qquad i = 1, \ldots, r, \qquad (21)$$

has a nontrivial real solution $\alpha_1, \ldots, \alpha_q$. Then $2 \leq q \leq k$, $\alpha_q \neq 0$, and the ratios $\alpha_1/\alpha_q, \ldots, \alpha_{q-1}/\alpha_q$ are uniquely determined, since otherwise the differences of the respective ratios would provide a nontrivial solution of (20). If we can show that these ratios are rational numbers, the theorem will result by taking a suitable common integral multiple of the numbers

$$x_m = \begin{cases} \dfrac{\alpha_m}{\alpha_q} & \text{for } 1 \leq m \leq q, \\[2mm] 0 & \text{for } q < m \leq k. \end{cases}$$

If we put $\alpha_m/\alpha_q = -\beta_m$ for $m = 1, \ldots, q-1$, equations (21) imply that

$$\log |\epsilon_q^{(i)}| = \sum_{m=1}^{q-1} \beta_m \log |\epsilon_m^{(i)}|, \qquad i = 1, \ldots, n. \qquad (22)$$

Now consider the set of all units η with the property that

$$\log |\eta^{(i)}| = \sum_{m=1}^{q-1} \gamma_m \log |\epsilon_m^{(i)}|, \qquad i = 1, \ldots, n, \qquad (23)$$

for suitable real numbers $\gamma_1, \ldots, \gamma_{q-1}$. For such an η the coefficients γ_m are unique. We call the set $\gamma_1, \ldots, \gamma_{q-1}$ of real numbers *proper* if η, as defined in (23) with these γ's, actually is a unit, and if in addition $|\gamma_1| < 1, \ldots, |\gamma_{q-1}| < 1$. If $\gamma_1, \ldots, \gamma_{q-1}$ is a proper set, then

$$|\log|\eta^{(i)}|| < \sum_{m=1}^{q-1} |\log|\epsilon_m^{(i)}||,$$

and, by Theorem 2-40, there are only finitely many (say H) proper

sets. On the other hand, if $\gamma_1, \ldots, \gamma_{q-1}$ is proper, so also is

$$N\gamma_1 - [N\gamma_1], \ldots, N\gamma_{q-1} - [N\gamma_{q-1}],$$

if N is a rational integer. For

$$\sum_{m=1}^{q-1} (N\gamma_m - [N\gamma_m]) \log |\epsilon_m^{(i)}| = \log |\eta^{(i)}|^N - \sum_{m=1}^{q-1} \log |\epsilon_m^{(i)}|^{[N\gamma_m]},$$

which is the logarithm of a product of powers of units, and is there-
fore the logarithm of a unit. Now if any β_m were irrational, then no
two of the numbers $N\beta_m - [N\beta_m]$, where N runs over Z, would be
equal, and we should have infinitely many proper sets. This con-
tradiction establishes the theorem.

THEOREM 2–44. *If $\epsilon_1, \ldots, \epsilon_k$ are units such that the only real
solution of (18) is the trivial solution, then there is a rational integer
M with the following property: in order that a number η such that*

$$\log |\eta^{(i)}| = \sum_{m=1}^{k} \gamma_m \log |\epsilon_m^{(i)}|, \qquad i = 1, \ldots, n,$$

*be a unit of K, it is necessary that all the numbers $M\gamma_m$ be rational
integers.*

Proof: The hypothesis is that which was used in the preceding
proof, except that we have replaced $q - 1$ by k. Suppose that
$\gamma_m = a/b$, where a and b are rational integers with $b > 0$ and $(a, b) = 1$,
and m is one of the integers $1, \ldots, k$. Then $N\gamma_m - [N\gamma_m]$ assumes
the b values $0/b, 1/b, \ldots, (b - 1)/b$, so that $b \leq H$, where H is the
number of proper sets. Hence, $b|H!$, and we can take $M = H!$.

THEOREM 2–45. *The group E of all units of K has a finite basis, the
number of basis elements of infinite order being at most r.*

Proof: The system (18) of r linear homogeneous equations in k
unknowns is certainly nontrivially solvable in reals if $k > r$, and it
follows from Theorem 2–43 that there are at most r independent
units in K. Let k be the exact maximal number of independent units,
and let $\epsilon_1, \ldots, \epsilon_k$ be such a set. Then by Theorem 2–44, for every
unit η of K there are $g_1, \ldots, g_k$ in Z such that

$$\log |\eta^{(i)}| = \sum_{m=1}^{k} \frac{g_m}{M} \log |\epsilon_m^{(i)}|, \qquad i = 1, \ldots n.$$

By the second part of Theorem 2–40, and Theorem 2–41, it follows that

$$\eta^M = \epsilon_1{}^{g_1} \cdots \epsilon_k{}^{g_k} \zeta_0{}^g,$$

so that $\epsilon_1, \ldots, \epsilon_k, \zeta_0$ form a basis for the group of Mth powers of units. Now define the numbers $\xi_0, \ldots, \xi_k$ by the equations

$$\xi_0 = \zeta_0{}^{1/M}, \qquad \xi_1 = \epsilon_1{}^{1/M}, \qquad \ldots, \qquad \xi_k = \epsilon_k{}^{1/M},$$

where an arbitrary but fixed Mth root is taken in each case. The numbers ξ_m may not lie in K, but they form a basis for a group E_M of complex numbers, and E_M clearly contains E as a subgroup. The theorem is therefore a consequence of the following general principle.

THEOREM 2–46. *If G is a commutative group having a basis of n elements, every subgroup of G also has a basis, of at most n elements.*

Proof: Suppose that $\lambda_1, \ldots, \lambda_n$ is a basis for G, that S is a subgroup of G, and that some λ_i actually occurs in the representation of some element s of S. Let I_i be the set of all exponents which occur on λ_i in the representations of the various elements of S. If a is in I_i, so is ka for k in Z, and if a and a' are in I_i, so is $a - a'$. Hence I_i is an ideal in Z, and is therefore a principal ideal, say $I_i = [a_i{}^*]$.

We now proceed by induction on n. If $n = 1$, then $\lambda_1{}^{a_1{}^*}$ is a basis for S, by what we have just proved. Suppose that the theorem is true for every commutative group with $n - 1$ basis elements, and suppose that G has n basis elements, say $\lambda_1, \ldots, \lambda_n$. Let S be a subgroup of G. If every element of S can be written in the form

$$\lambda_1{}^{a_1} \cdots \lambda_{n-1}{}^{a_{n-1}},$$

the result follows from the induction hypothesis. Otherwise, suppose that $I_n = [a]$, and let λ be an element of S in whose basis representation λ_n occurs with exponent a. Then for every s in S there exists a b_s in Z such that $s\lambda^{b_s}$ has a representation

$$s\lambda^{b_s} = \lambda_1{}^{a_1} \cdots \lambda_{n-1}{}^{a_{n-1}}.$$

The set of numbers of the form $s\lambda^{b_s}$ is therefore a subgroup of the group G' which has $\lambda_1, \ldots, \lambda_{n-1}$ as a basis, and by the induction hypothesis this subgroup also has a basis, of at most $n - 1$ elements. This latter basis, together with λ, clearly constitutes a basis for S.

REFERENCES

Section 2–4

The complete tabulation of Euclidean domains is the work of many writers. K. Inkeri (*Annales Academiae Scientiarum Fennicae, Series A* (Helsinki) I, Mathematics-Physics, **41**, 35pp. (1947)) supplied the last link in a chain of theorems which together show that if $d > 100$, then $R(\sqrt{d})$ is not Euclidean. E. S. Barnes and H. P. F. Swinnerton-Dyer (*Acta Mathematica* (Stockholm) **87**, 259–323 (1952)) showed that, contrary to what had been believed, $R(\sqrt{97})$ is not Euclidean. P. Varnavides (*Proceedings Konink. Nederlandsche Akademie van Wetenschappen, Series A* (Amsterdam) **55**, 111–122 (1952) or *Indagationes Mathematicae* (Amsterdam) **14**, 111–122 (1952)) showed that the values of d listed in the text yield Euclidean domains.

Section 2–9

The material of this section is adapted from E. Hecke, *Vorlesungen über die Theorie der Algebraischen Zahlen*, Leipzig: Akademische Verlagsgesellschaft m.b.H., 1923; reprinted by Chelsea Publishing Company, New York, 1948; pp. 116–131. It is proved there that the upper bound obtained in the text is exact.

CHAPTER 3

APPLICATIONS TO RATIONAL NUMBER THEORY

3–1 Introduction. As was suggested in the preceding chapter, there are many problems in rational number theory which are most naturally treated in the more extensive framework of an algebraic number field. Chief among these are various Diophantine equations; indeed, it was the study of Fermat's equation, $x^n + y^n = z^n$, $n \geq 3$, which was originally responsible for the development of ideal theory. While this approach has not led to a complete verification of Fermat's conjecture in all cases, it has produced results which would probably never have been obtained using rational methods alone. In the first part of this chapter we will discuss some results of this kind due to E. Kummer. Here heavy use will be made of ideal theory.

The latter portion of the chapter is primarily concerned with a theorem due to B. Delauney and T. Nagell, which asserts that the cubic analog of Pell's equation,

$$x^3 + dy^3 = 1,$$

has at most one solution in nonzero rational integers x, y, and completely characterizes this possible solution. (In the next chapter we shall prove a less precise result about the general equation $x^n + dy^n = 1$, $n \geq 3$.) Use is made here of the insolvability in Z of

$$x^3 + y^3 = z^3,$$

but otherwise the two parts are mutually independent.

3–2 Equivalence and class number. We say that the ideals A and B of $R[\vartheta]$ are *equivalent*, and write $A \sim B$, if there are nonzero elements α and β of $R[\vartheta]$ such that

$$[\alpha]A = [\beta]B.$$

It is easily seen that "$\sim$" is an equivalence relation. Moreover, if

$A \sim B$ and $C \sim D$, then $AC \sim BD$, and if $AC \sim BC$ then $A \sim B$.

THEOREM 3-1. *All principal ideals are equivalent. Any ideal equivalent to a principal ideal is principal.*

Proof: The first statement is trivial, since

$$[\alpha][\beta] = [\beta][\alpha].$$

If $A \sim [\alpha]$, then for some β and γ,

$$[\beta]A = [\alpha][\gamma] = [\alpha\gamma],$$

and hence

$$[\beta]|[\alpha\gamma],$$

$$\beta|\alpha\gamma,$$

$$\alpha\gamma = \beta\delta,$$

$$[\beta]A = [\alpha\gamma] = [\beta][\delta],$$

$$A = [\delta].$$

Since equivalence is an equivalence relation, the ideals of $R[\vartheta]$ can be separated into equivalence classes in the usual way. The number h of such classes is called the *class number* of the field; according to Theorem 3-1, $h = 1$ if and only if every ideal is principal, i.e., if and only if $R[\vartheta]$ is a unique factorization domain. We shall now show that h is always finite.

THEOREM 3-2. *There is a positive constant c, which depends only on the field, such that each ideal A divides a principal ideal AB for which*

$$\mathbf{N}AB \leq c\mathbf{N}A.$$

Proof: Let $\rho_1, \ldots, \rho_n$ be a field basis, and let $\rho_1^{(s)}, \ldots, \rho_n^{(s)}$ $(s = 1, \ldots, n)$ be the field conjugates of these numbers. We shall show that the theorem is true with

$$c = \prod_{s=1}^{n} (|\rho_1^{(s)}| + \cdots + |\rho_n^{(s)}|).$$

Let A be an arbitrary ideal, and let k be the greatest rational integer not exceeding $(\mathbf{N}A)^{1/n}$, so that $k^n \leq \mathbf{N}A < (k + 1)^n$. Then if $t_1, \ldots, t_n$ range independently over the integers $0, 1, \ldots, k$, there

are determined $(k + 1)^n$ different integers

$$t_1\rho_1 + \cdots + t_n\rho_n,$$

and two of them must be congruent modulo A:

$$u_1\rho_1 + \cdots + u_n\rho_n \equiv v_1\rho_1 + \cdots + v_n\rho_n \pmod{A}.$$

Thus

$$\alpha = (u_1 - v_1)\rho_1 + \cdots + (u_n - v_n)\rho_n$$

is in A, so that $A|[\alpha]$, and

$$\mathbf{N}[\alpha] = |\mathbf{N}\alpha| = \left| \prod_{s=1}^{n} \left(\sum_{i=1}^{n} (u_i - v_i)\rho_i^{(s)} \right) \right| \leq \prod_{s=1}^{n} \sum_{i=1}^{n} |u_i - v_i| \cdot |\rho_i^{(s)}|$$

$$\leq \prod_{s=1}^{n} \sum_{i=1}^{n} k|\rho_i^{(s)}| = ck^n \leq c\mathbf{N}A.$$

THEOREM 3–3.   *The class number of any algebraic number field is finite.*

Proof: It suffices to show that in each class there is an ideal B such that $\mathbf{N}B \leq c$, by the corollary to Theorem 2–35.   Let C be an arbitrary ideal of a given class, and determine A so that AC is principal. Then by Theorem 3–2, there is an ideal B such that AB is principal and $\mathbf{N}AB \leq c\mathbf{N}A$.   Then $AB \sim AC$, $B \sim C$, and

$$\mathbf{N}B = \frac{\mathbf{N}AB}{\mathbf{N}A} \leq c.$$

THEOREM 3–4.   *If h is the class number, the hth power of any ideal is principal.*

Proof: If $A_1, \ldots, A_h$ is a complete system of representatives of the various classes, and A is arbitrary, then $AA_1, \ldots, AA_h$ is another such system.   Hence

$$A_1 \cdots A_h \sim AA_1 \cdots AA_h = A^h A_1 \cdots A_h,$$

so $A^h \sim [1]$ and A^h is principal.

THEOREM 3–5.   *If p is a rational prime and $p \nmid h$, then $A^p \sim B^p$ implies $A \sim B$.*

Proof: Since $p \nmid h$, there are positive x and y in Z such that

$$px - hy = 1.$$

From the fact that $A^p \sim B^p$ we have

$$[\alpha]A^p = [\beta]B^p,$$

$$[\alpha]^x A^{px} = [\beta]^x B^{px},$$

$$[\alpha]^x A^{hy}A = [\beta]^x B^{hy}B,$$

and by Theorem 3–4, $A \sim B$.

Theorem 3–5 shows that the primes which do not divide h enjoy a property not shared by other primes. This is of great importance in the investigation of Fermat's equation.

<center>PROBLEMS</center>

1. Let α and β be algebraic integers, not both zero. Show that there is an integer δ such that, first, $\delta|\alpha$ and $\delta|\beta$ (in the sense that α/δ and β/δ are again integers), and, second, for suitable integers ξ and η, $\delta = \alpha\xi + \beta\eta$. Show that this GCD is unique up to an algebraic unit (i.e., an integer which divides 1). [*Hint:* First settle the case $\alpha\beta = 0$. In the other case, let K be an algebraic number field of class number h, containing both α and β. Then $[\alpha, \beta]^h = [\gamma]$, for some γ in K. Let δ be an integer such that $\delta^h = \gamma$, and show that the equation $[\alpha, \beta]^h = [\gamma]$ still holds when $[\alpha, \beta]$ and $[\gamma]$ are interpreted as ideals in $K(\delta)$. Deduce that $[\alpha, \beta] = [\delta]$ in $K(\delta)$.] Does the Unique Factorization Theorem hold in the domain of all algebraic integers?

2. Let K be an algebraic number field. Show that to each ideal A of K there corresponds an integer α (not necessarily in K) such that the elements of A are exactly those integers of K which are divisible by α.

3–3 The cyclotomic field K_p. Let p be an odd prime, let

$$\Phi(x) = \frac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + 1,$$

and let $\zeta = e^{2\pi i/p}$, so that the zeros of Φ are $\zeta, \zeta^2, \ldots, \zeta^{p-1}$, the primitive pth roots of unity. The field $R(\zeta) = R(\zeta^2) = \cdots = R(\zeta^{p-1}) = K_p$ is called a *cyclotomic field*. It is clearly of degree $p - 1$ at most. We put $1 - \zeta = \pi$. (The fact that the symbol π is used for two different numbers should occasion no confusion; the number $\pi = 3.14159 \ldots$ will occur only in the argument of the exponential function.)

THEOREM 3–6. *In K_p, the ideal $[p]$ has the factorization*

$$[p] = [\pi]^{p-1};$$

$[\pi]$ *is prime, and* $\mathbf{N}[\pi] = p;$ Φ *is irreducible, and* K_p *is of degree* $p - 1$.

Proof: Since ζ is an integer of K_p, so is

$$\epsilon_r = \frac{1 - \zeta^r}{1 - \zeta} = 1 + \zeta + \cdots + \zeta^{r-1}, \qquad 1 \le r \le p - 1;$$

if now an r' is chosen so that $rr' \equiv 1 \pmod{p}$, then $\zeta^{rr'} = \zeta$, so that

$$\epsilon_r^{-1} = \frac{1 - \zeta^{rr'}}{1 - \zeta^r} = 1 + \zeta^r + \cdots + \zeta^{r(r'-1)}$$

is also an integer.   Hence ϵ_r is a unit of K_p, and

$$p = \Phi(1) = \prod_{r=1}^{p-1} (1 - \zeta^r) = (1 - \zeta)^{p-1} \prod_{r=1}^{p-1} \epsilon_r = \epsilon(1 - \zeta)^{p-1},$$

where ϵ is a unit.   It follows from this equation that $[p] = [\pi]^{p-1}$, and also that $\mathbf{N}\pi = p$.   By Theorem 2–39, deg $K_p \ge p - 1$, so that $[\pi]$ is prime, deg $K_p = p - 1$, and Φ is irreducible.   (For a different proof of the irreducibility of Φ, see Problem 1, Section 2–2.)

Hereafter, we designate $[\pi]$ by P.

THEOREM 3–7.   *Writing* $\Delta(1, \zeta, \ldots, \zeta^{p-2}) = \Delta(\zeta),$ *we have*

$$\Delta(\zeta) = (-1)^{(p-1)/2} p^{p-2}.$$

Proof: From the representations

$$\Phi(x) = \prod_{r=1}^{p-1} (x - \zeta^r) = \frac{x^p - 1}{x - 1},$$

we obtain

$$\Phi'(\zeta^s) = \prod_{\substack{1 \le r \le p-1 \\ r \ne s}} (\zeta^s - \zeta^r) = \frac{p(\zeta^{rs} - 1)\zeta^{s(p-1)} - (\zeta^{ps} - 1)}{(\zeta^s - 1)^2}$$

$$= \frac{p\zeta^{s(p-1)}}{\zeta^s - 1} = -\frac{p}{\zeta^s(1 - \zeta^s)}.$$

Since

$$\Delta(\zeta) = \begin{vmatrix} 1 & \zeta & \cdots & \zeta^{p-2} \\ 1 & \zeta^2 & \cdots & \zeta^{2(p-2)} \\ \vdots & \vdots & & \vdots \\ 1 & \zeta^{p-1} & \cdots & \zeta^{(p-2)(p-1)} \end{vmatrix}^2 = \prod_{1 \le r < s \le p-1} (\zeta^r - \zeta^s)^2,$$

we have

$$\Delta(\zeta) = (-1)^{\frac{1}{2}(p-1)(p-2)} \prod_{s=1}^{p-1} \prod_{\substack{1 \leq r \leq p-1 \\ r \neq s}} (\zeta^s - \zeta^r)$$

$$= (-1)^{\frac{1}{2}(p-1)} \prod_{s=1}^{p-1} \Phi'(\zeta^s) = (-1)^{\frac{1}{2}(p-1)} \frac{(-1)^{p-1}p^{p-1}}{N\zeta N\pi}$$

$$= (-1)^{\frac{1}{2}(p-1)} p^{p-2}.$$

THEOREM 3-8. *The numbers* $1, \zeta, \ldots, \zeta^{p-2}$ *form an integral basis for* K_p, *so that*

$$\Delta = \Delta(\zeta) = (-1)^{\frac{1}{2}(p-1)} p^{p-2}.$$

Proof: Suppose that α is an integer of K_p, and that

$$\alpha = r_0 + r_1\zeta + \cdots + r_{p-2}\zeta^{p-2},$$

where the r's are rational. Then for $k = 0, 1, \ldots, p - 1$,

$$\zeta^k \alpha = \sum_{j=0}^{p-2} r_j \zeta^{j+k},$$

and since the trace function is clearly additive,

$$S(\zeta^k \alpha) = \sum_{j=0}^{p-2} S(r_j \zeta^{j+k}) = \sum_{j=0}^{p-2} r_j S(\zeta^j \zeta^k).$$

Solving this system of equations for the numbers r_j, we obtain

$$r_j = \frac{\text{a determinant in } \alpha \text{ and } \zeta}{\det |S(\zeta^j \zeta^k)|}.$$

But as we saw in the proof of Theorem 2-38, $\det |S(\zeta^j \zeta^k)| = \Delta(\zeta)$; since the determinant in the numerator has the rational value $r_j \Delta(\zeta)$, and is clearly an integer of K_p, it is a rational integer. Thus α can be written in the form

$$\alpha = \frac{c_0 + c_1\zeta + \cdots + c_{p-2}\zeta^{p-2}}{p^{p-2}} = \frac{d_0 + d_1\pi + \cdots + d_{p-2}\pi^{p-2}}{p^{p-2}},$$

where the c's, and therefore also the d's, are in Z. Since α is an integer,

$$p | (d_0 + d_1\pi + \cdots + d_{p-2}\pi^{p-2}),$$

and since $P|[p]$,

$$P|[d_0 + d_1\pi + \cdots + d_{p-2}\pi^{p-2}],$$

so $P|[d_0]$. It follows that $\mathbf{N}P|\mathbf{N}[d_0]$, $p|d_0^{p-1}$, and finally $p|d_0$. This argument may be repeated $p - 2$ times, to show that $p|d_k$ for $k = 1, \ldots, p - 2$, so that

$$\alpha = \frac{e_0 + e_1\pi + \cdots + e_{p-2}\pi^{p-2}}{p^{p-3}},$$

where the e's are rational integers. Repeating the entire argument $p - 3$ times, we see that

$$\alpha = f_0 + f_1\pi + \cdots + f_{p-2}\pi^{p-2},$$

where the f's are in Z. Hence $1, \pi, \ldots, \pi^{p-2}$ form an integral basis for K_p. But from the equations

$$\pi = 1 - \zeta, \qquad\qquad \zeta = 1 - \pi,$$

$$\pi^2 = 1 - 2\zeta + \zeta^2, \quad\text{and}\quad \zeta^2 = 1 - 2\pi + \pi^2,$$
$$\vdots \qquad\qquad\qquad \vdots$$

we see that $\Delta(\pi) = a^2\Delta(\zeta)$ and $\Delta(\zeta) = a^2\Delta(\pi)$, where a is a certain determinant with binomial coefficients as entries. Hence $a^2 = 1$, $\Delta(\zeta) = \Delta(\pi)$, and $1, \zeta, \ldots, \zeta^{p-2}$ also form an integral basis, by Theorem 2–14.

THEOREM 3–9.   *If α is an integer of K_p, there is a rational integer, a, such that*

$$\alpha^p \equiv a \pmod{P^p}.$$

Proof: Since $\mathbf{N}P = p$, the incongruent numbers $0, 1, \ldots, p - 1$ form a complete residue system modulo P, so that for suitable b in Z,

$$\alpha \equiv b \pmod{P}.$$

But

$$\alpha^p - b^p = \prod_{r=0}^{p-1} (\alpha - \zeta^r b),$$

and since $\zeta \equiv 1 \pmod{P}$,

$$\alpha^p - b^p \equiv \prod_{r=0}^{p-1} (\alpha - b) \equiv 0 \pmod{P^p},$$

so that we can take $a = b^p$.

If $P{\nmid}[\alpha]$ and $\alpha \equiv a \pmod{P^2}$ for some a in Z, then α is said to be *primary*.

THEOREM 3–10.   *If $\pi \nmid \alpha$, then for some positive rational integer f, $\zeta^f \alpha$ is primary.*

Proof: For suitable a and b in Z,

$$\alpha \equiv a + b\pi \pmod{P^2},$$

and $\pi \nmid \alpha$, so that $p \nmid a$. Choose f so that

$$af \equiv b \pmod{p}.$$

Then since

$$\zeta^f = (1 - \pi)^f \equiv 1 - f\pi \pmod{P^2},$$

we have

$$\zeta^f \alpha \equiv (1 - \pi f)(a + b\pi) \equiv a + \pi(-af + b) \equiv a \pmod{P^2}.$$

We now investigate the units of K_p.

THEOREM 3–11.   *The only roots of unity in K_p are the numbers $\pm\zeta^r$, $0 \le r < p$.*

Proof: The roots of unity are the numbers

$$e^{2\pi i t/m},$$

where t and m are rational integers, and $(t, m) = 1$. If such a number is in K_p, and if $tt' \equiv 1 \pmod{m}$, then also

$$e^{2\pi i tt'/m} = e^{2\pi i/m}$$

is in K_p. The numbers mentioned in the theorem are the $(2p)$th roots of unity, so we need only show that $e^{2\pi i/m}$ is not in K_p if $m \nmid 2p$. If $m \nmid 2p$, then either $4 \mid m$, or some odd prime $q \ne p$ divides m, or $p^2 \mid m$. Suppose that $e^{2\pi i/m}$ is in K_p.

If $4 \mid m$, then

$$e^{2\pi i/4} = i$$

is in K_p. But then so are $1 + i$ and $1 - i$, and

$$[1 + i] = [1 - i] \qquad \text{and} \qquad [2] = [1 + i]^2,$$

contrary to Theorem 2–38.

If $q \mid m$, then

$$\sigma = e^{2\pi i/q}$$

is in K_p. But then the reasoning used in the proof of Theorem 3–6

shows that

$$[q] = [1 - \sigma]^{q-1},$$

again contradicting Theorem 2–38.

If $p^2 | m$, then

$$\xi = e^{2\pi i/p^2}$$

is in K_p. But ξ is a zero of

$$\frac{x^{p^2} - 1}{x^p - 1} = x^{p(p-1)} + \cdots + x^p + 1 = \prod_{\substack{1 \le m \le p^2 \\ p \nmid m}} (x - \xi^m),$$

and

$$p = \prod_{\substack{1 \le m \le p^2 \\ p \nmid m}} (1 - \xi^m).$$

As before, the factors in this product are associated, and we get

$$[p] = [1 - \xi]^{p(p-1)},$$

contradicting Theorem 2–39.

THEOREM 3–12.   *Each unit ϵ of K_p can be written in the form*

$$\epsilon = \zeta^g \eta,$$

where g is a positive rational integer and η is real.

Proof: Express ϵ in terms of the integral basis $1, \zeta, \ldots, \zeta^{p-2}$:

$$\epsilon = f(\zeta),$$

where f is a polynomial with rational integral coefficients. Then clearly $\epsilon_s = f(\zeta^s)$ is also a unit, since $\mathbf{N}\epsilon = \epsilon_1 \cdots \epsilon_{p-1} = \pm 1$. Also,

$$\epsilon_{p-s} = f(\zeta^{p-s}) = f(\zeta^{-s}) = \overline{f(\zeta^s)} = \overline{\epsilon_s},$$

where the bar denotes the complex conjugate, so that $\epsilon_s \epsilon_{p-s} = |\epsilon_s|^2 > 0$, and

$$\mathbf{N}\epsilon = \prod_{s=1}^{\frac{1}{2}(p-1)} \epsilon_s \epsilon_{p-s} > 0,$$

so that $\mathbf{N}\epsilon = 1$.

Since $\epsilon_s = \overline{\epsilon_{p-s}}$,

$$\left| \frac{\epsilon_s}{\epsilon_{p-s}} \right| = 1.$$

The polynomial

$$\prod_{s=1}^{p-1} \left(x - \frac{\epsilon_s}{\epsilon_{p-s}} \right) = \prod_{s=1}^{p-1} (\epsilon_{p-s} x - \epsilon_s)$$

has coefficients in Z, so, by Theorem 2–40, $\epsilon_1/\epsilon_{p-1}$ is a root of unity, and by Theorem 3–11,

$$\epsilon_1 = \pm \zeta^m \epsilon_{p-1}.$$

Since either m or $p + m$ is even, and since

$$\zeta^m = \zeta^{p+m},$$

we can write

$$\epsilon_1 = \pm \zeta^{2g} \epsilon_{p-1}.$$

The proof will be complete if it can be shown that the plus sign is appropriate here, since then the quantities $\epsilon_1 \zeta^{-g}$ and $\epsilon_{p-1} \zeta^g$ are simultaneously equal and complex-conjugate, and are therefore real, so that $\epsilon = \epsilon_1 = \zeta^g (\epsilon_{p-1} \zeta^g)$.

To show this, choose a from among $0, 1, \ldots, p - 1$ so that

$$\zeta^{-g} \epsilon \equiv a \pmod{P}.$$

Then

$$\mu = \frac{\zeta^{-g} \epsilon - a}{\pi}$$

is an integer in K_p, as is

$$\bar{\mu} = \frac{\bar{\zeta}^{-g} \bar{\epsilon} - a}{\bar{\pi}} = \frac{\zeta^g \epsilon_{p-1} - a}{\bar{\pi}}.$$

Since $\bar{\pi} = 1 - \zeta^{p-1}$ is an associate of π, it follows that

$$\frac{\zeta^g \epsilon_{p-1} - a}{\pi}$$

is an integer of K_p, so that

$$\zeta^g \epsilon_{p-1} \equiv a \equiv \zeta^{-g} \epsilon \pmod{P}.$$

Thus

$$\frac{\epsilon}{\epsilon_{p-1}} = \pm \zeta^{2g} \qquad \text{and} \qquad \frac{\epsilon}{\epsilon_{p-1}} \equiv \zeta^{2g} \pmod{P}.$$

If the minus sign obtains in the equation, we have

$$-\zeta^{2g} \equiv \zeta^{2g} \pmod{P},$$

$$2\zeta^{2g} \equiv 0 \pmod{P},$$

$$P|[2\zeta^{2g}],$$

$$\mathbf{N}P|2^{p-1},$$

contrary to the fact that $\mathbf{N}P = p$. The proof is complete.

<div align="center">PROBLEMS</div>

1. Let p and q be distinct odd primes, and let ζ be a primitive pth root of unity.

   (a) Show that

   $$p = \prod_{a=1}^{p-1} (1 - \zeta^{2a}) = (-1)^{(p-1)/2} \prod_{a=1}^{\frac{1}{2}(p-1)} (\zeta^a - \zeta^{-a})^2.$$

   (b) Show that

   $$(\zeta^a - \zeta^{-a})^q \equiv \zeta^{qa} - \zeta^{-qa} \pmod{q}.$$

   (c) Deduce that

   $$p^{\frac{1}{2}(q-1)} \equiv (-1)^{\frac{1}{2}(p-1)\cdot\frac{1}{2}(q-1)} \prod_{a=1}^{\frac{1}{2}(p-1)} \frac{\zeta^{aq} - \zeta^{-aq}}{\zeta^a - \zeta^{-a}} \pmod{q}.$$

   (d) Show that the second factor on the right side of the last congruence above is $(-1)^\mu$, where μ is the number of numerically smallest residues $\pmod{p}$ among $q, 2q, \ldots, \frac{1}{2}(p - 1)q$ which are negative, and so obtain a proof of the law of quadratic reciprocity.

2. For an odd prime p and a positive integer h, put

   $$\Phi_h(x) = \frac{x^{p^h} - 1}{x^{p^{h-1}} - 1} = x^{p^{h-1}(p-1)} + x^{p^{h-1}(p-2)} + \cdots + x^{p^{h-1}} + 1,$$

and let ζ be a zero of Φ_h and $K_{p^h} = R(\zeta)$. Then the degree of K_{p^h} is at most $\varphi(p^h) = t$. Put $1 - \zeta = \pi$, and $[\pi] = P$.

   (a) Show that in K_{p^h}, the ideal $[p]$ has the factorization P^t, P is prime, $\mathbf{N}P = p$, $\Phi_h(x)$ is irreducible and K_{p^h} is of degree t.

   (b) Show that $\Delta(\zeta) = (-1)^{\frac{1}{2}t(t-1)} \cdot p^{p^{h-1}(hp-h-1)}$. [*Hint:* Notice that $1 - \zeta^{p^{h-1}}$ is a zero of $\Phi_1(1 - x)$, an irreducible polynomial of degree $p - 1$ with leading coefficient $(-1)^{p-1}$ and constant term p, and deduce that $\mathbf{N}(1 - \zeta^{p^{h-1}}) = ((-1)^{p-1}p)^{p^{h-1}}.$]

   (c) Show that if L is any prime ideal in K_{p^h} different from P, and if $\zeta^a \equiv \zeta^b \pmod{L}$, then $a \equiv b \pmod{p^h}$.

3–4 Fermat's equation. For the sake of completeness, we consider first the equation

$$x^n + y^n = z^n \tag{1}$$

for the cases $n = 2, 4$, and 3. When these have been disposed of, Fermat's assertion would be proved if it could be shown that (1) has no solutions in rational integers x, y, z, with $xyz \neq 0$, if n is a prime larger than 3.

The proof that (1) is impossible when $n = 4$ depends on the following theorem, which characterizes the solutions of (1) when $n = 2$.

THEOREM 3–13. *A general primitive solution* (*i.e., a solution in which* $(x, y, z) = 1$) *of*

$$x^2 + y^2 = z^2, \qquad y \text{ even}, \quad x > 0, \quad y > 0, \quad z > 0$$

is given by

$$x = a^2 - b^2, \qquad y = 2ab, \qquad z = a^2 + b^2,$$

where a *and* b *are prime to each other and not both odd, and* $a > b > 0$.

Remark: It is clear that one of x and y must be even, since otherwise $x^2 + y^2 \equiv 2 \not\equiv z^2$ (mod 4). There is no loss in generality in assuming that it is y which is even.

Proof: Suppose that $x^2 + y^2 = z^2$. Since $(x, y, z) = 1$, also $(y, z) = 1$, so that $(z - y, z + y) = 1$ or 2. But z is odd and y is even, so that $(z - y, z + y) = 1$. Hence, from the equation

$$x^2 = (z - y)(z + y),$$

we deduce that $z - y$ and $z + y$ must be squares, since they are positive. Now if t and u are fixed integers of the same parity (both odd or both even), there are integers a and b such that $t = a + b$ and $u = a - b$. Hence we can put

$$z - y = (a - b)^2, \qquad z + y = (a + b)^2,$$

which gives

$$z = \frac{(a - b)^2 + (a + b)^2}{2} = a^2 + b^2,$$

$$y = \frac{(a + b)^2 - (a - b)^2}{2} = 2ab,$$

$$x = (a - b)(a + b) = a^2 - b^2.$$

Since $(z - x, z + x) = (2a^2, 2b^2) = 2$, we must choose a and b so that $(a, b) = 1$. Since x is odd, $a + b$ must be odd. Since $y > 0$, a and b must have the same sign, and since $x > 0$, $|a| > |b|$. Since the pairs a, b and $-a, -b$ give the same solution, we can suppose that $a > b > 0$.

THEOREM 3–14.   *The equation $x^4 + y^4 = z^4$ is not solvable in non-zero rational integers.*

Proof: It suffices to show that there is no primitive solution of the equation

$$x^4 + y^4 = z^2.$$

Suppose that x, y, and z constitute such a solution; with no loss in generality we can take $x > 0, y > 0, z > 0$, and y even. Writing the supposed relation in the form

$$(x^2)^2 + (y^2)^2 = z^2,$$

we have from the preceding theorem that

$$x^2 = a^2 - b^2, \qquad y^2 = 2ab, \qquad z = a^2 + b^2,$$

where $(a, b) = 1$ and exactly one of a and b is odd. If a were even, we would have

$$1 \equiv x^2 = a^2 - b^2 \equiv -1 \pmod 4,$$

so $2|b$. We apply Theorem 3–13 again, this time to the equation $x^2 + b^2 = a^2$, and obtain

$$x = p^2 - q^2, \qquad b = 2pq, \qquad a = p^2 + q^2,$$

where $(p, q) = 1$, $p > q > 0$, and not both of p and q are odd. From

$$y^2 = 2ab$$

we have

$$y^2 = 4pq(p^2 + q^2).$$

Here p, q and $p^2 + q^2$ are relatively prime in pairs, so each must be a square:

$$p = r^2, \qquad q = s^2, \qquad p^2 + q^2 = t^2,$$

from which

$$r^4 + s^4 = t^2.$$

Now

$$x = r^4 - s^4, \qquad y = 2rst, \qquad z = a^2 + b^2 = r^8 + 6r^4s^4 + s^8,$$

so that $$z > (r^4 + s^4)^2 = t^4,$$

or $t < z^{\frac{1}{4}}$. It follows that if one solution of $x^4 + y^4 = z^2$ were known, another solution r, s, t could be found for which $rst \neq 0$ and $0 < t < z^{\frac{1}{4}}$. But this would give an infinite decreasing sequence of positive integers.

The case $n = 3$ is rather more difficult, since it is necessary to work in the quadratic field $K_3 = R(\zeta)$, where $\zeta = (-1 + i\sqrt{3})/2$ is a primitive cube root of unity. Not all the complications of the general case are present, however, since there is unique factorization of the integers of K_3, as the following theorem shows.

THEOREM 3–15. *Given any two integers α and γ of K_3, of which $\gamma \neq 0$, there are integers $\varkappa$ and ρ such that*

$$\alpha = \varkappa\gamma + \rho, \qquad 0 \leq \mathbf{N}\rho < \mathbf{N}\gamma.$$

The integers of K_3 therefore form a Euclidean domain.

Proof: Since 1 and ζ form an integral basis for K_3, we can write

$$\frac{\alpha}{\gamma} = \frac{a + b\zeta}{c + d\zeta} = \frac{(a + b\zeta)(c + d\zeta^2)}{c^2 - cd + d^2} = R + S\zeta,$$

where $a, b, c,$ and d are rational integers, and R and S are rational. Choose x and y in Z such that

$$|R - x| \leq \tfrac{1}{2}, \qquad |S - y| \leq \tfrac{1}{2};$$

then

$$\left|\frac{\alpha}{\gamma} - (x + y\zeta)\right|^2 = (R - x)^2 - (R - x)(S - y) + (S - y)^2 \leq \frac{3}{4}.$$

Hence, if $\varkappa = x + y\zeta$ and $\rho = \alpha - \varkappa\gamma$, then

$$\mathbf{N}\rho \leq \tfrac{3}{4}\mathbf{N}\gamma < \mathbf{N}\gamma, \qquad \text{and} \qquad \mathbf{N}\rho = \rho\bar{\rho} = |\rho|^2 \geq 0.$$

THEOREM 3–16. *The equation*

$$\xi^3 + \eta^3 + \vartheta^3 = 0 \tag{2}$$

has no solution in nonzero integers of K_3. It therefore has no solution in nonzero rational integers.

Proof: We first note that one of $\xi, \eta,$ and ϑ must be divisible by the prime $\pi = 1 - \zeta$, if (2) holds. For put

$$\xi + \eta = \rho, \qquad \eta + \vartheta = \sigma, \qquad \vartheta + \xi = \tau.$$

Then a simple calculation, using (2), shows that

$$(\rho + \sigma + \tau)^3 = 24\rho\sigma\tau.$$

Since the expression on the right side of this equation is divisible by

$$3 = -\zeta^2\pi^2,$$

the left side must be divisible by π, and therefore by π^3. Returning to the right side, it follows that one of ρ, σ, or τ must be divisible by π. If $\pi|\rho$, then $\pi|(\xi^3 + \eta^3)$, so $\pi|\vartheta^3$, and finally $\pi|\vartheta$.

If there were a common factor in two of ξ, η, and ϑ, it would also occur in the third, and could be divided out; so suppose that (2) holds, that ξ, η, and ϑ are relatively prime in pairs, and that $\pi|\vartheta$.

By Theorem 3–10, we may suppose that an appropriate power of ζ has been introduced into ξ and η so that

$$\xi \equiv 1, \qquad \eta \equiv -1 \ (\mathrm{mod}\ 3),$$

which we express by putting

$$\xi = 1 + 3\alpha, \qquad \eta = -1 + 3\beta,$$

where α and β are integers of K_3. Put

$$A = \frac{\xi + \zeta\eta}{\pi}, \qquad B = \frac{\zeta\xi + \eta}{\pi}, \qquad C = \frac{\zeta^2(\xi + \eta)}{\pi};$$

these numbers are integers of K_3, since

$$A = 1 + \frac{3}{\pi}(\alpha + \beta\zeta),$$

$$B = -1 + \frac{3}{\pi}(\zeta\alpha + \beta),$$

$$C = \frac{3}{\pi}(\alpha + \beta)\zeta^2.$$

Moreover,

$$A + B + C = 0, \tag{3}$$

$$ABC = \frac{\xi^3 + \eta^3}{\pi^3} = \left(\frac{-\vartheta}{\pi}\right)^3, \tag{4}$$

$$\xi = -\zeta A + \zeta^2 B, \qquad \eta = \zeta^2 A - \zeta B. \tag{5}$$

From (5) we see that $(A, B) = 1$, since otherwise ξ and η would have a common factor. From (3), also $(A, C) = (B, C) = 1$.

It follows from (4) that A, B, and C must all be cubes, say $A = \varphi^3$, $B = \chi^3$, $C = \psi^3$, and

$$\varphi^3 + \chi^3 + \psi^3 = 0.$$

Now $$A \equiv 1, \quad B \equiv -1, \quad C \equiv 0 \pmod{\pi},$$

so that from (4), ψ contains a smaller power of π than does ϑ.

Repeating the argument a sufficient number of times, we would arrive eventually at a solution of (2) in which no variable is divisible by π, which is impossible.

3-5 Kummer's theorem. If p is an odd rational prime, and its associated cyclotomic field K_p has class number h, then p is said to be *regular* if $p \nmid h$. According to Theorem 3-5, if p is a regular prime and A and B are ideals in K_p such that $A^p \sim B^p$, then $A \sim B$. It was this essential property of the regular primes which enabled Kummer to prove that Fermat's conjecture is correct for all regular primes. (Unfortunately, there are infinitely many irregular ones.) We shall not be able to prove Kummer's theorem in its entirety, but shall have to assume without proof a difficult preliminary result. We can, however, prove the following theorem.

THEOREM 3-17. *If p is regular, the equation*

$$x^p + y^p + z^p = 0 \tag{6}$$

has no solution in rational integers x, y, z for which $p \nmid xyz$.

Proof: Suppose that the theorem is false, and that x, y, and z satisfy all the requirements. We can assume that $(x, y) = 1$ and $p > 3$; as usual, ζ is a primitive pth root of unity, and $P = [1 - \zeta]$. From (6) we obtain

$$\prod_{m=0}^{p-1} (x + \zeta^m y) = -z^p,$$

so that

$$\prod_{m=0}^{p-1} [x + \zeta^m y] = [z]^p. \tag{7}$$

Now no two of the factors on the left have a common factor. For, if Q is a prime ideal such that $Q|[x + \zeta^{m_1}y]$ and $Q|[x + \zeta^{m_2}y]$ for $m_1 < m_2$, then

$$Q|[\zeta^{m_1}(1 - \zeta^{m_2-m_1})y],$$

and hence $Q|P[y]$. But from (7), $Q|[z]$, so $Q \neq P$ (since $p \nmid z$); hence $Q|[y]$. But then also $Q|[x]$, and we deduce that $NQ|y^{p-1}$ and $NQ|x^{p-1}$, which is contrary to the assumption that $(x, y) = 1$.

It follows that each factor on the left side of (7) is the pth power of an ideal. If

$$[x + \zeta y] = A^p,$$

then $A^p \sim [1] = [1]^p$, so that by Theorem 3-5, A itself is principal, say $A = [\alpha]$. Then

$$[x + \zeta y] = [\alpha]^p = [\alpha^p].$$

Hence

$$x + \zeta y = \epsilon \alpha^p,$$

where ϵ is a unit of K_p. Using the canonical form for units in K_p obtained in Theorem 3-12, we have

$$x + \zeta y = \zeta^g \eta \alpha^p, \qquad 0 \leq g \leq p - 1,$$

where η is real. By Theorem 3-9, since $[p]|P^p$,

$$\alpha^p \equiv a \pmod{[p]}$$

for some a in Z, so that

$$x + \zeta y \equiv \zeta^g \sigma \pmod{[p]},$$

where σ is a real integer of K_p. The complex conjugate of the integer

$$\frac{\zeta^{-g}(x + \zeta y) - \sigma}{p}$$

is also a field conjugate, and is therefore also an integer. Since $\bar{p} = p$ and $\bar{\sigma} = \sigma$, we have

$$\sigma \equiv (x + \zeta y)\zeta^{-g} \pmod{[p]},$$

and

$$\sigma \equiv (x + \zeta^{-1}y)\zeta^g \pmod{[p]},$$

so that

$$x\zeta^{-g} + y\zeta^{1-g} - x\zeta^g - y\zeta^{g-1} \equiv 0 \pmod{[p]}. \tag{8}$$

Two of these exponents must be congruent modulo p. For suppose that they are all distinct, and put

$$\beta = \frac{x}{p}\zeta^{-g} + \frac{y}{p}\zeta^{1-g} - \frac{x}{p}\zeta^g - \frac{y}{p}\zeta^{g-1}.$$

Then $p\beta$ has a representation in terms of distinct elements of an integral basis, the coefficients not being divisible by p. But since β is an integer, $p\beta$ also has a representation in which the coefficients are divisible by p, and this is contrary to the definition of a basis. We conclude that g must have one of the values 0, 1, or $(p + 1)/2$ (that is, $2g \equiv 1 \pmod{p}$).

If $g = 0$, the congruence (8) gives

$$y\zeta - y\zeta^{-1} \equiv 0 \pmod{[p]},$$

whence, since $\zeta^2 - 1$ is an associate of π,

$$P^2|[y]P, \qquad P|[y], \qquad p|y,$$

which is false. If $g = 1$, then (8) yields

$$x(1 - \zeta^2) \equiv 0 \pmod{[p]},$$

which implies that $p|x$, which is also false. Finally, if $g = (p + 1)/2$, then from (8) we get

$$(x - y)\pi \equiv 0 \pmod{[p]},$$

which gives

$$x \equiv y \pmod{p}.$$

Interchanging y and z in equation (7), we deduce that also

$$x \equiv z \pmod{p}.$$

But then equation (6) implies that

$$x^p + y^p + z^p \equiv 3x^p \equiv 0 \pmod{p},$$

which is false since $p > 3$ and $p \nmid x$. Hence the theorem is not false.

Because of its methodological interest, we deduce the general Kummer theorem from the following lemma, whose proof is too long for inclusion here:

KUMMER'S LEMMA. *Let p be a regular prime. Then if ϵ is a unit of K_p and a is a rational integer such that*

$$\epsilon \equiv a \pmod{P^p},$$

then ϵ is the pth power of another unit of K_p.

This is a partial converse of Theorem 3–9. Using it, we can generalize Theorem 3–17 in two ways: by allowing x, y, and z to be

integers of K_p instead of rational integers, or by dropping the restriction that $p \nmid xyz$.

THEOREM 3–18. *If p is a regular prime, the equation*

$$x^p + y^p + z^p = 0$$

has no solution in nonzero integers x, y, z of K_p for which $\pi \mid xyz$. It therefore has no solution in nonzero rational integers x, y, z for which $p \mid xyz$, and therefore (by Theorem 3–17) no nonzero rational integral solutions.

Proof: We first show that the equation

$$x^p + y^p = \epsilon \pi^{up} z'^{p}, \qquad \pi \nmid xyz', \quad \epsilon \text{ a unit of } K_p, \tag{9}$$

has no nontrivial solution if $u = 1$. Equation (9) is a generalization of the equation obtained from (6) by supposing that $z = z' \pi^u$, where $\pi \nmid z'$.

We may suppose that x and y have no common numerical factor, since it would also occur in z and could be canceled out. (Notice that it cannot be assumed that the ideals $[x]$ and $[y]$ are relatively prime, since $[x, y]$ may not be principal.) We may also suppose that x and y are primary, since they may be multiplied by appropriate powers of ζ without affecting (9). If (9) is written in the form

$$\prod_{m=0}^{p-1} (x + \zeta^m y) = \epsilon \pi^{up} z'^{p}, \tag{9'}$$

it is clear that at least one of the factors on the left, say $x + \zeta^i y$, is divisible by π. Since, however, the differences

$$(x + \zeta^k y) - (x + \zeta^i y) = (\zeta^k - \zeta^i)y$$

and

$$\zeta^i(x + \zeta^k y) - \zeta^k(x + \zeta^i y) = -(\zeta^k - \zeta^i)x$$

are also divisible by π, each factor on the left in (9') must be divisible by π. If two factors were divisible by π^2, we would have

$$\pi^2 \mid (\zeta^k - \zeta^i)y,$$

$$\pi^2 \mid \epsilon' \pi y \qquad (\epsilon' \text{ a unit}),$$

$$\pi \mid y,$$

and similarly $\pi \mid x$, contrary to assumption. On the other hand, since

x and y are primary, there is an a in Z such that

$$x + y \equiv a \pmod{P^2}.$$

But then

$$a \equiv x + y \equiv 0 \pmod{P},$$

$$p | a,$$

$$P^2 | [a],$$

$$x + y \equiv 0 \pmod{P^2}.$$

Thus the total number of factors of π on the left side of (9′) is at least $p + 1$, so that $u > 1$.

Now rewrite (9′) as

$$\prod_{m=0}^{p-1} [x + \zeta^m y] = P^{up}[z']^p. \tag{9″}$$

Any common factor different from P of two ideals in the product on the left side of (9″) must be a factor of both $[x]$ and $[y]$, and therefore of $[z']$. After dividing out every such common factor, as well as one factor of P from each ideal, the ideals remaining on the left are pairwise prime, and their product is a pth power; therefore each factor separately is a pth power.

Combining all these results, we can write

$$[x + y] = P^{p(u-1)+1} J_0{}^p \, D,$$

$$[x + \zeta^m y] = P \, J_m{}^p \, D, \qquad m = 1, \ldots, p - 1,$$

where $D = [x, y]$, and $J_0, J_1, \ldots, J_{p-1}$ are certain ideals not divisible by P. If we put $t_m = p(u - 1) + 1$ or 1, according as $m = 0$ or $m > 0$, we have, for $m \neq 1$,

$$[x + \zeta y] P^{t_m} J_m{}^p \, D = [x + \zeta y][x + \zeta^m y] = P J_1{}^p \, D[x + \zeta^m y];$$

since P is a principal ideal, it follows that

$$J_m{}^p \, D \sim J_1{}^p \, D,$$

so that $J_m{}^p \sim J_1{}^p$, and by Theorem 3–5, $J_m \sim J_1$. Thus integers γ_m and δ_m (which are not divisible by π) exist such that

$$[\gamma_m] J_m = [\delta_m] J_1, \qquad m = 0, 2, 3, \ldots, p - 1. \tag{10}$$

Raising both sides of (10) (with $m = 0$) to the pth power, and then multiplying through by $D \, P^{p(u-1)+1}$, we have

$$[\gamma_0]^p D \, P^{p(u-1)+1} \, J_0^p = D \, P \, J_1^p \cdot P^{p(u-1)}[\delta_0]^p,$$

or

$$[\gamma_0^p][x + y] = [x + \zeta y] \, P^{p(u-1)}[\delta_0^p].$$

Similarly,

$$D \, P[\gamma_2]^p J_2^p = D \, P[\delta_2]^p \, J_1^p,$$

$$[x + \zeta^2 y][\gamma_2^p] = [x + \zeta y][\delta_2^p],$$

so that

$$\gamma_0^p(x + y) = \epsilon_1(x + \zeta y)\pi^{p(u-1)}\delta_0^p,$$

$$\gamma_2^p(x + \zeta^2 y) = \epsilon_2(x + \zeta y)\delta_2^p, \tag{11}$$

where ϵ_1 and ϵ_2 are units.

We now use the identity

$$(x + \zeta^2 y) + (x + y)\zeta = (x + \zeta y)(1 + \zeta).$$

We multiply through by $\gamma_0^p \gamma_2^p$, and in the resulting equation replace the left sides of equations (11) by the right sides. After canceling the common factor $x + \zeta y$, there results

$$\epsilon_2(\gamma_0 \delta_2)^p + \epsilon_1 \zeta \pi^{p(u-1)}(\gamma_2 \delta_0)^p = (1 + \zeta)(\gamma_0 \gamma_2)^p.$$

Since ϵ_1, ϵ_2, ζ, and $1 + \zeta = (1 - \zeta^2)/(1 - \zeta)$ are units, this equation is of the form

$$\xi^p + \epsilon_3 \eta^p = \epsilon_4 \pi^{p(u-1)}\vartheta^p, \tag{12}$$

where ϵ_3 and ϵ_4 are units and $\pi \nmid \xi\eta\vartheta$. By Theorem 3–9,

$$\xi^p \equiv a_1, \qquad \eta^p \equiv a_2 \pmod{P^p},$$

where a_1 and a_2 are rational integers; since $u > 1$, (12) gives

$$a_1 + \epsilon_3 a_2 \equiv 0 \pmod{P^p}.$$

Since $\pi \nmid \eta$, also $\pi \nmid a_2$, so that $p \nmid a_2$. Choose a_3 so that $a_2 a_3 \equiv 1 \pmod{p^2}$; then

$$a_2 a_3 \equiv 1 \pmod{P^p},$$

$$a_1 a_3 + \epsilon_3 \equiv 0 \pmod{P^p}.$$

By Kummer's lemma, $\epsilon_3 = \epsilon_5^p$, and (12) becomes

$$\xi^p + (\epsilon_5 \eta)^p = \epsilon_4 \pi^{p(u-1)}\vartheta^p,$$

which is an equation of the form (9) with u replaced by $u - 1$. Repeating the argument $u - 2$ times, we would have a solution of (9) with $u = 1$, which is impossible.

Before leaving the subject of Fermat's conjecture, it might be of some interest to mention certain other facts known about it. We consider only the solvability of equation (6),

$$x^p + y^p + z^p = 0,$$

in Z.

It was proved by Wieferich in 1909 that if (6) holds in integers x, y, and z such that $p \nmid xyz$ (the so-called Case I), then

$$2^{p-1} \equiv 1 \pmod{p^2}.$$

Later investigators have shown that in Case I,

$$q^{p-1} \equiv 1 \pmod{p^2}$$

for every prime $q \leq 43$; J. B. Rosser used this fact to show that there are no solutions in Case I for $p < 41{,}000{,}000$. D. H. and Emma Lehmer later extended Rosser's method to prove Fermat's conjecture in Case I for $p < 253{,}747{,}889$. This in turn implies that if there is a solution in Case I, it must be that $\log \log z > 23$.

Without the restriction to Case I, Theorem 3–18 disposes of the regular primes. Kummer also found criteria to handle the irregular primes less than 164; this was pushed on to all $p < 619$ by H. S. Vandiver and his collaborators, and quite recently D. H. and E. Lehmer and Vandiver have used high-speed computing techniques to settle the problem for all $p < 2000$. It turns out that of the 302 primes less than 2000, 118 are irregular; while it is not known that there are infinitely many regular primes, there is nothing in the limited data available to indicate that there are only finitely many.

3–6 The equation $x^2 + 2 = y^3$. For the remainder of this chapter we shall be primarily concerned with the cubic analog of Pell's equation. At one point in the argument, however, we shall need the following auxiliary result.

THEOREM 3–19.   *The only solutions in Z of the equation*

$$x^2 + 2 = y^3 \tag{13}$$

are $x = \pm 5, y = 3$.

Proof: Following Euler's idea, we make use of the arithmetic of the quadratic field $R(\sqrt{-2})$. By Theorem 2–16, the integers of this

field are of the form $a + b\sqrt{-2}$, where a and b are rational integers. By a proof exactly paralleling that of Theorem 3–15, it can be shown that they form a Euclidean domain: given a, b, c, d in Z, with $cd \neq 0$, there are e, f, g, h in Z such that

$$a + b\sqrt{-2} = (c + d\sqrt{-2})(e + f\sqrt{-2}) + (g + h\sqrt{-2}),$$
$$g^2 + 2h^2 < c^2 + 2d^2.$$

It follows that $R[\sqrt{-2}]$ is a unique factorization domain.

We first show that if x and y satisfy (13), then $x + \sqrt{-2}$ and $x - \sqrt{-2}$ are relatively prime. It is clear that

$$(x + \sqrt{-2}, x - \sqrt{-2}) \mid -2\sqrt{-2},$$

and since $-2\sqrt{-2} = (\sqrt{-2})^3$ and $\sqrt{-2}$ is prime in the domain (by Theorem 2–39), it must be that $(x + \sqrt{-2}, x - \sqrt{-2}) = (\sqrt{-2})^m$, $0 \leq m \leq 3$. But if $x + \sqrt{-2} = (a + b\sqrt{-2})\sqrt{-2}$, then $x = -2b$, whence, by (13),

$$4b^2 + 2 = y^3,$$

$$y^3 \equiv 2 \pmod{4},$$

which is impossible.

Since the only units of $R(\sqrt{-2})$ are ± 1, it follows from (13) that

$$x + \sqrt{-2} = (a + b\sqrt{-2})^3,$$

where a and b are rational integers, and equating real and imaginary parts gives

$$a^3 - 6ab^2 = x,$$

$$3a^2b - 2b^3 = 1.$$

From the second of these equations it follows that $b = \pm 1$, and hence that $3a^2 - 2 = \pm 1$, or $a = \pm 1$. From the first, $x = \pm 1 \mp 6 = \pm 5$.

3–7 Pure cubic fields. The field $L = R(\sqrt[3]{d})$, in which $d > 1$ is a cube-free rational integer and $\sqrt[3]{d}$ is real, is called a *pure cubic field*. In this section we determine an integral basis for L and note certain other properties.

Since d is cube-free, we can write

$$d = ab^2,$$

where ab is square-free. Since $\sqrt[3]{d^2} = b\sqrt[3]{a^2b}$, the numbers 1, $\sqrt[3]{ab^2}$, $\sqrt[3]{a^2b}$ form a basis for L. Following Dedekind, we say that L is of the *first* or *second kind*, according as 9 does not, or does, divide $a^2 - b^2$. The reason for the distinction is made clear in the following theorem.

THEOREM 3-20. *The numbers*

$$1, \sqrt[3]{ab^2}, \sqrt[3]{a^2b}$$

form an integral basis for L if it is of the first kind. The numbers

$$\tfrac{1}{3}(1 + a\sqrt[3]{ab^2} + b\sqrt[3]{a^2b}), \sqrt[3]{ab^2}, \sqrt[3]{a^2b}$$

form an integral basis for L if it is of the second kind.

Remark: Note that the second basis represents every integer represented by the first, since

$$3z_1 \frac{1 + a\sqrt[3]{ab^2} + b\sqrt[3]{a^2b}}{3} + (z_2 - az_1)\sqrt[3]{ab^2} + (z_3 - bz_1)\sqrt[3]{a^2b}$$
$$= z_1 + z_2\sqrt[3]{ab^2} + z_3\sqrt[3]{a^2b}.$$

Proof: Suppose that ω is an integer in L, and that

$$\omega = x_1 + x_2\sqrt[3]{ab^2} + x_3\sqrt[3]{a^2b}, \qquad x_1, x_2, x_3 \text{ in } R.$$

Then the conjugates of ω are

$$\omega' = x_1 + \rho x_2\sqrt[3]{ab^2} + \rho^2 x_3\sqrt[3]{a^2b},$$
$$\omega'' = x_1 + \rho^2 x_2\sqrt[3]{ab^2} + \rho x_3\sqrt[3]{a^2b},$$

where ρ is a primitive cube root of unity. We see that

$$\omega + \omega' + \omega'' = 3x_1,$$
$$\sqrt[3]{a^2b}(\omega + \rho^2\omega' + \rho\omega'') = 3abx_2,$$
$$\sqrt[3]{ab^2}(\omega + \rho\omega' + \rho^2\omega'') = 3abx_3,$$

and since the left sides of these equations are algebraic integers and the right sides are rational, it follows that the numbers $3x_1$, $3abx_2$, $3abx_3$ are rational integers. Hence for any integer ω in L, there are y_1, y_2, y_3 in Z such that

$$3ab\omega = y_1 + y_2\sqrt[3]{ab^2} + y_3\sqrt[3]{a^2b}. \tag{14}$$

We show first that ab is a divisor of y_1, y_2, and y_3, and so can be omitted in (14).

Let p be a rational prime dividing a, and let P be a prime ideal of L which divides $[p]$. It was supposed that ab is square-free; a fortiori, $(a, b) = 1$, and $P \nmid [b]$. If we put

$$\alpha = \sqrt[3]{ab^2}, \qquad \beta = \sqrt[3]{a^2 b},$$

then $P|[\alpha]^3$, so $P^3|[\alpha]^3$; since L is of degree 3, it follows from Theorem 2–39 that $[p] = P^3$. Hence $P\||[\alpha]$ and $P^2\||[\beta]$.

Now suppose, in accordance with (14), that

$$y_1 + y_2\alpha + y_3\beta \equiv 0 \pmod{3ab}.$$

Then

$$y_1 + y_2\alpha + y_3\beta \equiv 0 \pmod{P^3},$$

$$y_1 \equiv 0 \pmod{P},$$

$$y_1 \equiv 0 \pmod{p}, \tag{15}$$

$$y_1 \equiv 0 \pmod{P^3},$$

$$y_2\alpha + y_3\beta \equiv 0 \pmod{P^3},$$

$$y_2\alpha \equiv 0 \pmod{P^2},$$

$$y_2 \equiv 0 \pmod{p}, \tag{16}$$

$$y_3\beta \equiv 0 \pmod{P^3},$$

$$y_3 \equiv 0 \pmod{p}. \tag{17}$$

By equations (15), (16), and (17), and the fact that p was an arbitrary prime divisor of a, we see that a divides y_1, y_2, and y_3. Similarly, b divides y_1, y_2, and y_3. It follows that there are z_1, z_2, z_3 in Z such that

$$3\omega = z_1 + z_2\alpha + z_3\beta. \tag{18}$$

Let the defining equation of ω be

$$x^3 + c_1 x^2 + c_2 x + c_3 = 0, \qquad c_1, c_2, c_3 \text{ in } Z.$$

Then by (18) and the analogous equations for $3\omega'$ and $3\omega''$,

$$c_1 = -(\omega + \omega' + \omega'') = -z_1,$$

$$c_2 = \omega\omega' + \omega\omega'' + \omega'\omega'' = \tfrac{1}{3}(z_1^2 - abz_2z_3), \tag{19}$$

$$c_3 = -\omega\omega'\omega'' = -\tfrac{1}{27}(z_1^3 + ab^2 z_2^3 + a^2 b z_3^3 - 3abz_1z_2z_3). \tag{20}$$

Suppose that $3|a$; then $3\nmid b$, and L is of the first kind. Since c_2 is in Z, $3|z_1$. Since c_3 is in Z,

$$0 \equiv -27c_3 \equiv 3\frac{a}{3}b^2 z_2{}^3 \ (\mathrm{mod}\ 9),$$

whence $3|z_2$, and by (20) again, $3|z_3$. In this case, then, the numbers 1, $\sqrt[3]{ab^2}$, $\sqrt[3]{a^2b}$ constitute an integral basis for L. A similar argument applies in the case that $3|b$.

Suppose now that $3\nmid ab$, so that

$$a^2 \equiv b^2 \equiv 1 \ (\mathrm{mod}\ 3). \tag{21}$$

If $3|z_1$, then by (19), $3|z_2 z_3$; if $3|z_2$, say, then it follows from (20) that also $3|z_3$. Similarly, if $3|z_2$, then also $3|z_1$ and $3|z_3$. Hence 3 divides all or none of z_1, z_2, z_3; in the first case ω is of the form specified in the theorem.

We now examine the possibility that ω in (19) is an integer, but that $3\nmid z_1 z_2 z_3$. Then by (20), (21), and Fermat's theorem,

$$z_1{}^3 + ab^2 z_2{}^3 + a^2 b z_3{}^3 \equiv 0 \ (\mathrm{mod}\ 3),$$

$$z_1 + az_2 + bz_3 \equiv 0 \ (\mathrm{mod}\ 3),$$

$$z_1 \equiv az_2 \equiv bz_3 \ (\mathrm{mod}\ 3),$$

$$z_2 \equiv az_1, \qquad z_3 \equiv bz_1 \ (\mathrm{mod}\ 3),$$

$$z_2 = az_1 + 3t_2, \qquad z_3 = bz_1 + 3t_3.$$

Substituting these expressions for z_2 and z_3 into (20), we obtain

$$
\begin{aligned}
-27c_3 &= z_1{}^3 + ab^2(az_1 + 3t_2)^3 + a^2 b(bz_1 + 3t_3)^3 \\
&\quad - 3abz_1(az_1 + 3t_2)(bz_1 + 3t_3) \\
&= z_1{}^3(1 + a^4 b^2 + a^2 b^4 - 3a^2 b^2) \\
&\quad + 9z_1{}^2(a^3 b^2 t_2 + a^2 b^3 t_3 - a^2 b t_3 - ab^2 t_2) \\
&\quad + 27z_1(a^2 b^2 t_2{}^2 + a^2 b^2 t_3{}^2 - abt_2 t_3) + 27(ab^2 t_2{}^3 + a^2 b t_3{}^3) \\
&\equiv z_1{}^3(1 + a^4 b^2 + a^2 b^4 - 3a^2 b^2) \\
&\quad + 9z_1{}^2(ab^2 t(a^2 - 1) + a^2 b t_3(b^2 - 1)) \ (\mathrm{mod}\ 27).
\end{aligned}
$$

By (21),

$$0 \equiv -27c_3 \equiv z_1{}^3(1 + a^4 b^2 + a^2 b^4 - 3a^2 b^2) \ (\mathrm{mod}\ 27),$$

and it follows that

$$1 + a^4 b^2 + a^2 b^4 - 3a^2 b^2 \equiv 0 \ (\mathrm{mod}\ 27). \tag{22}$$

Using (21), we can put

$$b^2 = a^2 + 3f + 9g,$$

where f and g are rational integers and $0 \le f \le 2$. Then the congruence (22) reduces to

$$\varphi(f, a) = 2a^6 + (9f - 3)a^4 + 9(f^2 - f)a^2 + 1 \equiv 0 \pmod{27}.$$

For $f = 0$, this becomes

$$(a^2 - 1)^2(2a^2 + 1) \equiv 0 \pmod{27},$$

which is true for every a not divisible by 3, since

$$2a^2 + 1 \equiv a^2 - 1 \equiv 0 \pmod 3.$$

Moreover, for every a such that $3 \nmid a$,

$$\varphi(1, a) \equiv \varphi(0, a) + 9a^4 \equiv 9a^4 \not\equiv 0 \pmod{27},$$

$$\varphi(2, a) \equiv \varphi(0, a) + 18a^4 + 18a^2 \equiv 18a^2(a^2 + 1) \not\equiv 0 \pmod{27}.$$

Thus we find that if $3 \nmid z_1 z_2 z_3$, than c_3 is in Z if and only if

$$a^2 \equiv b^2 \not\equiv 0 \pmod 9,$$

(i.e., if and only if L is of the second kind) and $az_2 \equiv bz_3 \pmod 3$. If this is the case, then c_1 and c_2 are also rational integers, and

$$\omega = \tfrac{1}{3}(z_1 + (az_1 + 3t_2)\alpha + (bz_1 + 3t_3)\beta)$$

$$= z_1 \frac{1 + a\alpha + b\beta}{3} + t_2\alpha + t_3\beta$$

is an integer in L. The proof is complete.

In the course of the proof, it appeared that if $\omega = (x + y\alpha + z\beta)/3$ is an integer, and if one of x, y, z is divisible by 3, all of them are. In particular, if $x + y\alpha$ is an integer and x and y are rational, they are also integers.

We now consider the units of L. If L is of the first kind, then

$$\eta = x + y\alpha + z\beta$$

is a unit if and only if $\mathbf{N}\eta = \eta\eta'\eta'' = \pm 1$, or

$$x^3 + ab^2y^3 + a^2bz^3 - 3abxyz = \pm 1. \tag{23}$$

If L is of the second kind, then

$$\eta = \frac{u}{3}(1 + a\alpha + b\beta) + v\alpha + w\beta$$

is a unit if and only if

$$x^3 + ab^2 y^3 + a^2 b z^3 - 3abxyz = \pm 27, \qquad (24)$$

where $u = x$, $au + 3v = y$, and $bu + 3w = z$. If η is positive, the plus sign must be chosen in (23) and (24), since η' and η'' are complex conjugates.

The field L has the property that each of its elements is either rational or of degree three. For if there were an element of degree two, L would be an extension of the field generated by that element, and so would be of even degree. It follows that ± 1 are the only roots of unity in L. Since α has one real and two nonreal conjugates, we see by Theorem 2–45 that either L has only the units ± 1, or else there is a fundamental unit ξ, which may be chosen between 0 and 1, such that every unit η of L can be expressed in the form

$$\eta = \pm\xi^n,$$

where n is a rational integer, positive, negative, or zero.

A positive unit of the form $\eta = x + y\alpha$ is always smaller than 1. For since $x^3 + dy^3 = 1$, we have

$$\eta^{-1} = \eta'\eta'' = x^2 - xy\alpha + y^2\alpha^2 \geq 1 + \alpha + \alpha^2 > 3,$$

since xy is negative. Consequently, for such a unit we have

$$\eta = \xi^n, \qquad n > 0.$$

The same remarks apply to a positive unit of the form $x + z\beta$.

3–8 Two lemmas. For simplicity in notation, we define the binomial coefficient $\binom{m}{k}$ to be zero for $k > m$. Here and hereafter in this chapter, lower-case Latin letters stand for rational integers, unless otherwise specified.

THEOREM 3–21. *Let m be a positive integer. Then*

$$\binom{m}{0} + \binom{m}{3} + \binom{m}{6} + \cdots \not\equiv 0 \pmod{3}.$$

Proof: Put

$$S_0 = \binom{m}{0} + \binom{m}{3} + \binom{m}{6} + \cdots,$$

$$S_1 = \binom{m}{1} + \binom{m}{4} + \binom{m}{7} + \cdots,$$

$$S_2 = \binom{m}{2} + \binom{m}{5} + \binom{m}{8} + \cdots.$$

Then

$$S_0 + S_1 + S_2 = 2^m \equiv (-1)^m \,(\text{mod } 3),$$

and

$$S_2 = \binom{m}{1}\frac{m-1}{2} + \binom{m}{4}\frac{m-4}{5} + \cdots \equiv -mS_1 + S_1 \,(\text{mod } 3),$$

$$S_1 = \binom{m}{0}\frac{m}{1} + \binom{m}{3}\frac{m-3}{4} + \cdots \equiv mS_0 \,(\text{mod } 3),$$

so that

$$(1 + 2m - m^2)S_0 \equiv (-1)^m \,(\text{mod } 3).$$

THEOREM 3-22. *Suppose that x and y are integers such that $(x, dy) = 1$, and suppose that*

$$(x + y\sqrt[3]{d})^n = X + Y\sqrt[3]{d} + Z(\sqrt[3]{d})^2,$$

where X, Y, and Z are rational and $n > 1$. Then $XYZ \neq 0$ except in the following cases:

$$(\sqrt[3]{10} - 1)^5 = 99 - 45\sqrt[3]{10},$$

$$(\sqrt[3]{4} - 1)^4 = -15 + 12\sqrt[3]{2}.$$

Proof: Since $(x, d) = 1$, it is clear that $X \neq 0$. Suppose that $Z = 0$, so that

$$\binom{n}{2}x^{n-2}y^2 + \binom{n}{5}x^{n-5}y^5d + \binom{n}{8}x^{n-8}y^8d^2 + \cdots = 0. \quad (25)$$

Dividing by $\binom{n}{2}y^2$, this becomes

$$-x^{n-2} = \sum_{k \geq 1}\binom{n-2}{3k}\frac{2x^{n-3k-2}y^{3k}d^k}{(3k+1)(3k+2)}. \quad (26)$$

Let q be a prime divisor of y. Then since $q^{3k} \geq 2^{3k} > 3k + 2$ for $k \geq 1$, each term in the last sum is divisible by q, which is impossible since $(x, y) = 1$. Hence $y = \pm 1$.

When $n \equiv 0 \pmod 3$, equation (26) can be written in the form

$$-y^{n-3}d^{\frac{1}{3}(n-3)} = \sum_{k \geq 1} \binom{n-1}{3k} \frac{x^{3k}y^{n-3k-3}d^{\frac{1}{3}(n-3)-k}}{3k+1};$$

when $n \equiv 1 \pmod 3$,

$$-y^{n-4}d^{\frac{1}{3}(n-4)} = \sum_{k \geq 1} \binom{n-2}{3k} \frac{2x^{3k}y^{n-3k-4}d^{\frac{1}{3}(n-4)-k}}{(3k+1)(3k+2)};$$

and when $n \equiv 2 \pmod 3$,

$$-y^{n-2}d^{\frac{1}{3}(n-2)} = \sum_{k \geq 1} \binom{n}{3k} x^{3k}y^{n-3k-2}d^{\frac{1}{3}(n-2)-k}.$$

The same argument now shows that $x = \pm 1$, and since it is clear from (25) that $xy < 0$, we have $x = -y$.

Now let q be a prime divisor of d, and suppose that $q^{\alpha} \| d$ (that is, $q^{\alpha} | d$ but $q^{\alpha+1} \nmid d$). If $q^{\alpha} > 5$, then $q^{\alpha k} > 5^k \geq 3k + 2$ for $k \geq 1$, so that each term in the sum in (26) is divisible by q, which is impossible since $(x, d) = 1$. If $q = 3$, then since $3 \nmid (3k + 1)(3k + 2)$ we reach the same contradiction. Hence $q^{\alpha} = 2$ or 5, and $d = 2$, 5, or 10.

The information obtained so far shows that

$$1 - \binom{n-2}{3}\frac{2d}{45} + \binom{n-2}{6}\frac{2d^2}{7 \cdot 8} - \cdots = 0. \tag{27}$$

If $d = 10$, this becomes

$$\binom{n-2}{3} - 1 = \frac{1}{6}(n-5)(n^2 - 4n + 6)$$

$$= \sum_{k \geq 2} (-1)^k \binom{n-2}{3k} \frac{2 \cdot 10^k}{(3k+1)(3k+2)}.$$

This equation is true for $n = 5$, and leads to the first of the exceptions mentioned in the theorem. For other values of n, we may divide through by $(n-5)/6$ and obtain

$$n^2 - 4n + 6$$

$$= -\sum_{k \geq 1} (-1)^k \binom{n-6}{3k-1} \frac{(n-2)(n-3)(n-4) \cdot 12 \cdot 10^{k+1}}{3k(3k+1)(3k+2)(3k+3)(3k+4)(3k+5)}.$$

The highest power of 5 which divides the denominator of a term in the sum is clearly at most $5(3k + 5)$, and since $5^{k+1} > 5(3k + 5)$ for $k \geq 2$, we have

$$n^2 - 4n + 6 = (n-2)^2 + 2 \equiv \binom{n-6}{2} \frac{(n-2)(n-3)(n-4) \cdot 12 \cdot 10^2}{3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8} \equiv 0 \pmod 5,$$

which is false since -2 is a quadratic nonresidue of 5.

When $d = 2$ or 5, equation (27) leads to the congruence

$$1 + \binom{n-2}{3} + \binom{n-2}{6} + \cdots \equiv 0 \pmod 3,$$

which is false by Theorem 3–21.

There remains only the possibility that $Y = 0$. The proof that this happens only in the case of the second exception mentioned in the theorem is completely similar to what has just been done for the case $Z = 0$, and we leave the details to the reader. (The only variation lies in the fact that d may now have the sole prime divisor 2, so that $d = 2$ or 4.)

3–9 The Delaunay-Nagell theorem. As we shall see in the next chapter, there is a general theorem which implies that the equation

$$ax^3 + by^3 = c \tag{28}$$

has only finitely many solutions in integers x, y if a, b, and c are nonzero integers. In certain special cases, however, it is possible to make more precise statements about the number and nature of possible solutions. We shall concern ourselves here with the equation

$$x^3 + dy^3 = 1, \tag{29}$$

which was first considered in detail by B. Delaunay. His method was later refined by T. Nagell, who also applied it to (28) in the case that $c = 1$ or 3. Nagell's result concerning (29) is as follows.

THEOREM 3–23. *Equation (29) has at most one solution in integers x, y different from zero. If x_1, y_1 is a solution, the number $x_1 + y_1 \sqrt[3]{d}$ is either the fundamental unit of $L = R(\sqrt[3]{d})$ or its square; the latter can happen for only finitely many values of d.*

If $d = \pm 1$, (29) has only trivial solutions. If d contains a cube larger than 1, it can be absorbed into the factor y^3. Hence we can assume that d is cube-free and larger than 1.

The idea of the proof is quite simple. If

$$\mathbf{N}(x_1 + y_1 \sqrt[3]{d}) = x_1{}^3 + dy_1{}^3 = 1, \qquad y_1 \neq 0,$$

then $x_1 + y_1 \sqrt[3]{d}$ is a positive unit of L, and as such is a positive power of the fundamental unit ξ mentioned at the end of Section 3–7. It therefore suffices to show that no power of a positive unit smaller than 1, with exponent larger than 2, is of the special form $x + y\sqrt[3]{d}$, and to show that the square of a unit is of this form in only finitely many cases. We divide the proof into four parts, summarized in the next four theorems.

THEOREM 3–24.   *The square of an irrational unit of L of the form*

$$\eta = x + y\alpha + z\beta, \qquad x, y, z \text{ in } Z$$

is itself of the form $X + Y\alpha$ only if

$$\eta = 1 + \sqrt[3]{20} - \sqrt[3]{50}.$$

The square of a unit of L of the form

$$\eta = \tfrac{1}{3}(x + y\alpha + z\beta), \qquad 3 \nmid xyz,$$

(if such exists) is itself of the form $X + Y\alpha$ for only finitely many values of d.

Proof: Let $\qquad \eta = x + y\alpha + z\beta$

be a positive unit of L, so that, by (23),

$$x^3 + ab^2y^3 + a^2bz^3 - 3abxyz = 1 \qquad (30)$$

and

$$\eta^2 = (x^2 + 2abyz) + (2xy + az^2)\alpha + (2xz + by^2)\beta.$$

If the coefficient of β in this last expression is 0, then

$$z = -\frac{by^2}{2x},$$

and substituting this into (30) we obtain

$$x^3 + dy^3 - d^2 \frac{y^6}{8x^3} + 3d \frac{y^3}{2} = 1,$$

or $$d^2y^6 - 20x^3dy^3 - 8(x^6 - x^3) = 0,$$

whence $$dy^3 = 10x^3 \pm 2x\sqrt{27x^4 - 2x}. \tag{31}$$

Thus the number $27x^4 - 2x$ must be a square:

$$(27x^3 - 2)x = t^2. \tag{32}$$

If x is even, then $(27x^3 - 2, x) = 2$, so that

$$27x^3 - 2 = \pm 2u^2, \qquad x = \pm 2v^2.$$

Since -1 is a quadratic nonresidue of 3, we must choose the lower sign, and eliminating x we obtain

$$108v^6 + 1 = u^2,$$

$$(u - 1)(u + 1) = 108v^6.$$

Since $(u - 1, u + 1) = 2$, this implies that

$$u \pm 1 = 54r^6, \qquad u \mp 1 = 2s^6,$$

whence

$$27r^6 - s^6 = (3r^2)^3 - (s^2)^3 = \pm 1.$$

From the truth of Fermat's conjecture for $n = 3$, it follows that $r = 0$, which gives $v = 0$ and $x = 0$. But then also $y = 0$, by (31), which is impossible since $z\beta$ is not a unit.

If x is odd, (32) yields

$$27x^3 - 2 = \pm u^2, \qquad x = \pm v^2.$$

Here the upper sign must be chosen, and we have

$$(3x)^3 = u^2 + 2,$$

which by Theorem 3–19 has the sole solution $x = 1, u = \pm 5$. By (31), $dy^3 = 10 \pm 10$, so that $d = 20, y = 1$. (If $y = 0$, then $z = 0$, and η is rational.) The sole solution is therefore

$$(1 + \sqrt[3]{20} - \sqrt[3]{50})^2 = -19 + 7\sqrt[3]{20}.$$

Now let η be a positive unit of the form

$$\eta = \tfrac{1}{3}(x + y\alpha + z\beta).$$

Then by (24), $$x^3 + ab^2y^3 + a^2bz^3 - 3abxyz = 27, \tag{33}$$

and $$9\eta^2 = (x^2 + 2abyz) + (2xy + az^2)\alpha + (2xz + by^2)\beta. \tag{34}$$

If $3|x$, also $3|y$ and $3|z$, and we have already treated this case. Suppose that $3\nmid x$. If the coefficient of β in the expression for η^2 is 0, we again have

$$z = -\frac{by^2}{2x}.$$

Substituting this into (33), it follows that

$$dy^3 = 10x^3 \pm 6x\sqrt{3x^4 - 6x}, \tag{35}$$

so that

$$3x^4 - 6x = t^2. \tag{36}$$

If x is even, the fact that $3\nmid x$ implies that

$$x^3 - 2 = \pm 6u^2, \qquad x = \pm 2v^2,$$

whence

$$\pm 4v^6 - 1 = \pm 3u^2.$$

Since $3\nmid(4v^6 + 1)$, we must choose the upper sign; the last equation can then be written as

$$(u + 1)^3 - (u - 1)^3 = (2v^2)^3,$$

so that $|u| = 1$. Hence $x = 2$, and by (35), $dy^3 = 80 \pm 72$. The lower sign yields $d = 1$ or 8, both of which are excluded. Hence $d = 19$, $y = 2$, and $z = -1$. The only solution in this case is

$$\left(\frac{2 + 2\sqrt[3]{19} - (\sqrt[3]{19})^2}{3}\right)^2 = -8 + 3\sqrt[3]{19}.$$

If x is odd, (36) implies that

$$x^3 - 2 = \pm 3u^2, \qquad x = \pm v^2,$$

so that

$$\pm v^6 - 2 = \pm 3u^2.$$

The lower sign must be chosen: $x = -v^2$ and

$$3u^2 - 2 = v^6. \tag{37}$$

But it is an immediate consequence of Theorem 4–17, to be proved in the next chapter, that (37) has only finitely many solutions, and the proof is complete.

We note for future use that if u, v satisfy (37), then v must be odd.

THEOREM 3–25. *The fourth power of a positive irrational unit of L is never of the form $X + Y\alpha$.*

Proof: Let ϵ be such a unit,

$$\epsilon = \tfrac{1}{3}(x_1 + y_1\alpha + z_1\beta),$$

and suppose that

$$\epsilon^4 = X + Y\alpha.$$

Then since the coefficient of β in ϵ^4 is 0, we have

$$6bx_1{}^2y_1{}^2 + 4x_1{}^3z_1 + 4ab^2y_1{}^3z_1 + 12abx_1y_1z_1{}^2 + a^2bz_1{}^4 = 0. \quad (38)$$

If we put

$$\eta = \epsilon^2 = \tfrac{1}{3}(x + y\alpha + z\beta),$$

then

$$x = \tfrac{1}{3}(x_1{}^2 + 2aby_1z_1),$$
$$y = \tfrac{1}{3}(2x_1y_1 + az_1{}^2),$$
$$z = \tfrac{1}{3}(2x_1z_1 + by_1{}^2).$$

Since $\eta^2 = X + Y\alpha$, we can apply Theorem 3–24. The cases

$$d = 20, \qquad x = y = -z = 3,$$
$$d = 19, \qquad x = y = 2, \quad z = -1$$

are impossible, since in the first the above equation for z becomes $-9 = 2x_1z_1 + 2y_1{}^2$, while in the second the system is easily seen to be inconsistent for all choices of signs of x_1, y_1, z_1. Hence it must be that $x = -v^2$, where v is odd, so that

$$3v^2 + x_1{}^2 = -2aby_1z_1.$$

Since v is odd, so is x_1, so that $3v^2 + x_1{}^2 \equiv 4 \pmod 8$. Hence three of the numbers a, b, y_1, z_1 are odd, and the fourth is even. By (38), $a^2bz_1{}^4$ is even, so y_1 is odd. If either a or z_1 is even, (38) implies that $6bx_1{}^2y_1{}^2 \equiv 0 \pmod 4$, which is false. If b is even, (38) implies that $a^2bz_1{}^4 \equiv 0 \pmod 4$, which is false since b is square-free. The proof is complete.

THEOREM 3–26. *The cube of a positive irrational unit of L is never of the form $X + Y\alpha$.*

Proof: If

$$\eta = \tfrac{1}{3}(x + y\alpha + z\beta)$$

is a positive unit, the coefficient of β in η^3 is

$$\tfrac{1}{9}(bxy^2 + x^2z + abyz^2).$$

We see from the equation

$$x^3 + ab^2y^3 + a^2bz^3 - 3abxyz = 27 \tag{39}$$

that $(x, b) = 1$, and deduce from the equation

$$bxy^2 + x^2z + abyz^2 = 0 \tag{40}$$

that $b|z$. From (39) again, $\delta = (x, y, z) = 1$ or 3. Since $\eta \neq \pm 1$, y and z are not both zero, and we can write

$$x = \delta d_1 d_2 x_1, \qquad y = \delta d_2 d_3 y_1, \qquad z = \delta b d_1 d_3 z_1, \tag{41}$$

where

$$\left(\frac{x}{\delta}, \frac{z}{\delta b}\right) = |d_1|, \qquad \left(\frac{x}{\delta}, \frac{y}{\delta}\right) = |d_2|, \qquad \left(\frac{y}{\delta}, \frac{z}{\delta b}\right) = |d_3|,$$

and $x_1 > 0, y_1 > 0, z_1 > 0$. The numbers d_1x_1, d_2y_1, d_3z_1 are relatively prime in pairs. Substituting the values from (41) into (39) and dividing by $\delta^3 b d_1 d_2 d_3$, we obtain

$$d_2{}^2 d_3 x_1 y_1{}^2 + d_1{}^2 d_2 x_1{}^2 z_1 + ab^2 d_1 d_3{}^2 y_1 z_1{}^2 = 0.$$

It follows from this that $d_1|x_1, d_2|y_1$, and $d_3|z_1$. Putting

$$x_1 = d_1 x_2, \qquad y_1 = d_2 y_2, \qquad z_1 = d_3 z_2,$$

substituting, and dividing by $d_1 d_2 d_3$, we obtain

$$d_2{}^3 x_2 y_2{}^2 + d_1{}^3 x_2{}^2 z_2 + ab^2 d_3{}^3 y_2 z_2{}^2 = 0.$$

A consequence of this is that $x_2|ab^2 d_3{}^3 y_2 z_2{}^2$, which in turn implies that $x_2 = 1$. Similarly, $y_2 = z_2 = 1$, so that

$$d_1{}^3 + d_2{}^3 + ab^2 d_3{}^3 = 0 \tag{42}$$

and

$$x = \delta d_1{}^2 d_2, \qquad y = \delta d_2{}^2 d_3, \qquad z = \delta b d_1 d_3{}^2.$$

Substituting these values into (39), we obtain

$$d_1{}^6 d_2{}^3 + ab^2 d_2{}^6 d_3{}^2 + a^2 b^4 d_1{}^3 d_3{}^6 - 3ab^2 d_1{}^3 d_2{}^3 d_3{}^3 = \frac{27}{\delta^3}.$$

Eliminating $ab^2 d_3{}^3$ between this equation and (42), we have

$$d_1{}^9 + 6d_1{}^6 d_2{}^3 + 3d_1{}^3 d_2{}^6 - d_2{}^9 = \left(\frac{3}{\delta}\right)^3, \tag{43}$$

and putting $d_1{}^3 = u, d_2{}^3 = v, 3/\delta = w$, this becomes

$$u^3 + 6u^2v + 3uv^2 - v^3 = w^3. \tag{44}$$

But it is easily verified that

$$(u^3 + 6u^2v + 3uv^2 - v^3)U^3 = V^3 + W^3,$$

where

$$U = u^2 + uv + v^2, \qquad V = u^3 + 3u^2v - v^3, \qquad W = 3u^2v + 3uv^2.$$

Since neither U nor V is zero for relatively prime u and v, (44) can hold with $w \neq 0$ only if $W = 0$, that is, if $u = -v$. In this case $w = v$. Since $(d_1, d_2) = 1$, it follows that $d_1 = -1$, $d_2 = 1$, $\delta = 3$. This, however, leads to the values $x = 3$, $y = -3$, $z = 0$, for which the coefficient of β in η^3 is not zero.

THEOREM 3-27. *If $p > 3$ is prime and*

$$\eta = \tfrac{1}{3}(x + y\alpha + z\beta)$$

is a positive unit smaller than 1, then η^p is not of the form $X + Y\alpha$.

Proof: Suppose that $z = 0$. Then $3|x$ and $3|y$, and

$$\mathbf{N}\eta = \left(\frac{x}{3}\right)^3 + d\left(\frac{y}{3}\right)^3 = 1,$$

so that $\left(\dfrac{x}{3}, d\dfrac{y}{3}\right) = 1$; by Theorem 3-22, the coefficient of β in η^p is not zero. Thus $z \neq 0$. By the same reasoning (applied in the field $L' = R(\beta) = R(\alpha) = L$), y cannot be zero.

As we saw in the proof of Theorem 3-20, it follows from the representation

$$\omega = x_1 + x_2\alpha + x_3\beta$$

of an arbitrary integer ω of L that

$$\alpha(\omega + \rho\omega' + \rho^2\omega'') = 3abx_3.$$

Taking $\omega = \eta^p$, we see that if the coefficient of β in η^p is zero, it must be that

$$\left(\frac{x + y\alpha + z\beta}{3}\right)^p + \rho\left(\frac{x + y\rho\alpha + z\rho^2\beta}{3}\right)^p$$

$$+ \rho^2\left(\frac{x + y\rho^2\alpha + z\rho\beta}{3}\right)^p = 0. \qquad (45)$$

Suppose first that $p \equiv 1 \pmod 3$. Then since $\rho^p = \rho$, (45) can be written in the form

$$\left(\frac{x\rho + y\rho^2\alpha + z\beta}{3}\right)^p + \left(\frac{x\rho^2 + y\rho\alpha + z\beta}{3}\right)^p = -\left(\frac{x + y\alpha + z\beta}{3}\right)^p.$$

Since p is odd, the left side is divisible by

$$\frac{x\rho + y\rho^2\alpha + z\beta}{3} + \frac{x\rho^2 + y\rho\alpha + z\beta}{3} = \frac{-x - y\alpha + 2z\beta}{3};$$

this number is an integer, and since it divides η^p, it is a unit. Consequently,

$$-x^3 - ab^2y^3 + 8a^2bz^3 - 6abxyz = \pm 27.$$

Since η is a positive unit, also

$$x^3 + ab^2y^3 + a^2bz^3 - 3abxyz = 27,$$

and by addition,

$$9a^2bz^3 - 9abxyz = 0 \text{ or } 54.$$

In the first case $az^2 - xy$ must be zero. But this number is the coefficient of α in $3/\eta$, and as we saw at the end of Section 3-7, it is not zero, since $1/\eta > 1$.

In the second case we have

$$abz(az^2 - xy) = 6.$$

But then x, y, z are not all divisible by 3, so that L is of the second kind. This is impossible, since if $ab|6$ then $a^2 - b^2 \not\equiv 0 \pmod 9$.

The case in which $p \equiv 2 \pmod 3$ proceeds similarly. Equation (45) can be written in the form

$$\left(\frac{x\rho^2 + y\alpha + z\rho\beta}{3}\right)^p + \left(\frac{x\rho + y\alpha + z\rho^2\beta}{3}\right)^p = -\left(\frac{x + y\alpha + z\beta}{3}\right)^p,$$

from which it follows that the number

$$\frac{x\rho^2 + y\alpha + z\rho\beta}{3} + \frac{x\rho + y\alpha + z\rho^2\beta}{3} = \frac{-x + 2y\alpha - z\beta}{3}$$

is a unit. As before,

$$9ab^2y^3 - 9abxyz = 0 \text{ or } 54.$$

Since $by^2 - xz$ is the coefficient of β in $3/\eta$, it is not zero. But it is also impossible that $(by^2 - xz)|6$ and $ab|6$, since then L must be of both the first and second kinds. The proof is complete.

Theorems 3-25, 3-26, and 3-27 show that any nonzero solution of $x^3 + dy^3 = 1$ must correspond either to the fundamental unit of L, or to its square. Not both of these numbers can lead to solutions, by Theorem 3-22 with $n = 1$. This completes the proof of Theorem 3-23.

REFERENCES

Section 3-5

For a complete exposition of what is known concerning Fermat's conjecture, see H. S. Vandiver, "Fermat's last theorem: the history and the nature of the results concerning it," *American Mathematical Monthly* **53**, 555–578 (1946). The result of Lehmer, Lehmer, and Vandiver was announced in *Proceedings of the National Academy of Sciences* **40**, 25–33 (1954). Landau gives a proof of Kummer's lemma; see his *Vorlesungen über Zahlentheorie*, vol. 3, Leipzig: S. Hirzel Verlag, 1927.

Section 3-6

The equation $y^2 = x^3 + k$ was the subject of L. J. Mordell's inaugural address, *A Chapter in the Theory of Numbers*, New York: Cambridge University Press, 1947. Also see Dickson's *History of the Theory of Numbers*, Washington: Carnegie Institution of Washington, 1919; reprinted, Chelsea Publishing Company, New York, 1950; vol. 2, pp. 531–539.

Section 3-7

Dedekind's fundamental paper on pure cubic fields is in *Journal für die Reine und Angewandte Mathematik* (Berlin) **121**, 40–123 (1899).

Sections 3-8, 3-9

We have followed the treatment by Nagell, *Journal des Mathématiques Pures et Appliquées* (Paris) **4**, 209–270 (1925). Delaunay (*Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* (Paris) **171**, 336 (1920) and **172**, 434 (1921)) announced that equation (28) has at most five solutions in case $c = 1$. His work on (29) was announced in *Comptes Rendus* **162**, 150–151 (1916).

CHAPTER 4

THE THUE-SIEGEL-ROTH THEOREM

4–1 Introduction. It is shown in introductory texts in number theory* that if α is a quadratic irrationality (that is, an algebraic number of degree two), then there is a positive constant c such that

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^2}$$

for every pair of rational integers p, q with $q > 0$. The idea used there suffices to prove the following generalization, which is due to J. Liouville.

THEOREM 4–1. *If α is an algebraic number of degree $n \geq 2$, then there exists a positive constant c such that*

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^n} \tag{1}$$

for every pair of rational integers p, q with $q > 0$.

Proof: Let α be a zero of the irreducible polynomial

$$f(x) = a_0 x^n + \cdots + a_n, \qquad a_0 > 0,$$

with coefficients in Z, and let $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n$ be its conjugates, so that

$$f(x) = a_0(x - \alpha)(x - \alpha_2) \cdots (x - \alpha_n).$$

Then the number

$$q^n f\left(\frac{p}{q}\right) = a_0 p^n + a_1 p^{n-1} q + \cdots + a_n q^n$$

is a rational integer different from zero, and it therefore has absolute

---

* See for example, Volume I, Section 8–4. In Section 8–5 Hurwitz' theorem is stated and proved, and in Chapter 9 the problem of approximating real numbers by rationals is considered; all this material is assumed in the present section.

121

value at least 1. Hence

$$\left| \alpha - \frac{p}{q} \right| = \frac{\left| q^n f\left(\frac{p}{q}\right) \right|}{a_0 q^n \prod_{k=2}^{n} \left| \alpha_k - \frac{p}{q} \right|} \geq \frac{1}{a_0 q^n \prod_{k=2}^{n} \left| \alpha_k - \frac{p}{q} \right|}.$$

Put

$$\beta = \overline{|\alpha|} = \max(|\alpha|, \ldots, |\alpha_n|).$$

We consider two cases, according as $|p/q|$ is greater than 2β or not. In the first case we have the trivial lower bound

$$\left| \alpha - \frac{p}{q} \right| > \beta \geq \frac{\beta}{q^n}.$$

In the second case the inequality $|\alpha_k - p/q| \leq 3\beta$ holds for $k = 2, \ldots, n$, and, by the inequality of the preceding paragraph,

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{1}{a_0 q^n (3\beta)^{n-1}}.$$

Thus, the theorem holds with

$$c = \min\left(\beta, \frac{1}{a_0 (3\beta)^{n-1}} \right).$$

Liouville used this theorem to show the existence of nonalgebraic numbers; this will be discussed in detail in the next chapter. At the moment, let us consider a hypothetical improvement of Theorem 4–1, in which the inequality (1) is replaced by

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^\nu}, \tag{2}$$

where ν is any number smaller than n. A. Thue noticed that if such a theorem could be proved, it would have the important consequence that the Diophantine equation

$$q^n f\left(\frac{p}{q}\right) = a_0 p^n + a_1 p^{n-1} q + \cdots + a_n q^n = A \tag{3}$$

can have only finitely many solutions for any fixed rational integer A different from zero, if $f(x)$ has distinct zeros. To see this, let the zeros of $f(x)$ again be $\alpha_1 = \alpha, \ldots, \alpha_n$, and put

$$\gamma = \min_{i \neq j} (|\alpha_i - \alpha_j|).$$

Suppose that (3) has infinitely many solutions p, q. Then there must be at least one α_i, which by suitable naming we can take to be α, which is a limit point of the numbers p/q, since otherwise the quantity

$$q^n f\left(\frac{p}{q}\right) = a_0 q^n \prod_{k=1}^{n} \left(\frac{p}{q} - \alpha_k\right)$$

is certainly not bounded as q increases indefinitely. There must therefore be infinitely many solutions of (3) for which $|\alpha - p/q| < \gamma/2$. But for all such solutions,

$$\left|\alpha - \frac{p}{q}\right| = \frac{A}{a_0 q^n \prod\limits_{k=2}^{n} \left|\frac{p}{q} - \alpha_k\right|} \leq \frac{A}{a_0 \left(\frac{\gamma}{2}\right)^{n-1}} \cdot \frac{1}{q^n},$$

and this is at variance with (2) if ν is a constant smaller than n and q is sufficiently large.

Thue showed that (2) holds with

$$\nu = \frac{n}{2} + 1.$$

Later C. L. Siegel improved Thue's result, showing that (2) holds with

$$\nu > \min_{\substack{1 \leq s \leq n-1 \\ s \in Z}} \left(\frac{n}{s+1} + s\right),$$

and in particular with $\nu = 2\sqrt{n}$. In 1947 F. J. Dyson made the further improvement $\nu > \sqrt{2n}$, and finally in 1955 K. F. Roth proved that (2) holds with $\nu = 2 + \epsilon$, for each $\epsilon > 0$, for all but a finite number of fractions p/q. This is the best theorem possible if ν is to be independent of q, since Hurwitz' theorem shows that the corresponding statement is false for every irrational algebraic number, for $\nu = 2$ and suitable c. Roth's work is similar in some respects to a simplification of Dyson's proof, published by T. Schneider in 1948.

In addition to the problem of sharpening Theorem 4–1 by decreasing the exponent of q, we may also consider the question of extending the methods so as to analyze the approximability of an algebraic number by other algebraic numbers. This is not mere generalization for its own sake: as we saw in the preceding chapter, it is natural to

consider the solvability of $x^p + y^p = z^p$ in a larger set of integers than Z, and the same is true of many other Diophantine equations. But if the variables in an equation range over the integers of an algebraic number field, then to the extent that approximation theorems are useful at all they must be formulated in terms of algebraic rather than rational numbers.

While Siegel gave many algebraic variants of his basic result, Roth presented a detailed proof only in the rational case. In this chapter we give a complete proof of a useful algebraic version of Roth's theorem. Unfortunately, the proof is complicated; the student might profit by first examining Schneider's work mentioned above.

We shall proceed as follows. In the next three sections we shall make some definitions, and obtain some preliminary results, which are needed for the proof of the main theorem: in Section 4–2 some properties of polynomials will be treated, in Section 4–3 the concept of the generalized Wronskian will be introduced, and in Section 4–4 the index of a polynomial will be defined and discussed. Then we shall proceed to prove, in Sections 4–5 and 4–6, several lemmas on which the proof of the main theorem depends, and finally, in Section 4–7, we shall state and prove the Thue-Siegel-Roth theorem itself. In the remainder of the chapter, some applications of the theorem will be taken up.

4–2 Polynomials. If $P(z)$ is a polynomial with arbitrary complex coefficients, we denote by $||P||$ the maximum of the absolute values of its coefficients. If α is an algebraic number and $P(z) = 0$ is its defining equation, so that P is irreducible and has relatively prime coefficients in Z, we define the *height* $H(\alpha)$ to be $||P||$. Finally, if P has algebraic coefficients, we designate by $\overline{|P|}$ the maximum of the absolute values of their conjugates. Clearly $||P|| = \overline{|P|}$ if P has coefficients in Z, and for a nonzero constant polynomial $P(z) = \alpha$ the new definition of $\overline{|\alpha|}$ agrees with the old one.

Except when a polynomial is written as a determinant, it will be supposed that no two terms have the same exponents on the variable, or sets of exponents on the variables.

THEOREM 4–2. *Let $l, \lambda_1, \ldots, \lambda_h$ be complex numbers, and put*

$$L(z) = l \prod_{k=1}^{h} (z - \lambda_k).$$

Then
$$|l| \prod_{k=1}^{h} (1 + |\lambda_k|) \le 6^h ||L||.$$

Proof: There is no loss in generality in supposing that $l = 1$, since a change in l affects in the same way the two sides of the inequality to be proved. Let $\lambda_1, \ldots, \lambda_t$ be those of the λ's such that $|\lambda_k| \le 2$. If $f(z) = \prod_{k=1}^{t} (z - \lambda_k)$, then there is a complex number z_0 with $|z_0| = 1$ for which $|f(z_0)| \ge 1$. To see this, let ϵ be a $(t+1)$th root of unity, and suppose that

$$f(z) = \sum_{r=0}^{t} \mu_r z^r, \qquad \mu_t = 1.$$

Then
$$\sum_{\nu=0}^{t} \epsilon^\nu f(\epsilon^\nu) = \sum_{\nu=0}^{t} \epsilon^\nu \sum_{r=0}^{t} \mu_r \epsilon^{\nu r} = \sum_{r=0}^{t} \mu_r \sum_{\nu=0}^{t} \epsilon^{\nu(r+1)}.$$

But
$$\sum_{\nu=0}^{t} \epsilon^{\nu(r+1)} = \begin{cases} 0 & \text{if } (t+1) \nmid (r+1), \\ t+1 & \text{if } (t+1) | (r+1), \end{cases} \tag{4}$$

and since $r \le t$, $(t+1)|(r+1)$ if and only if $r = t$. Hence

$$\sum_{\nu=0}^{t} \epsilon^\nu f(\epsilon^\nu) = (t+1)\mu_t = t+1,$$

so that one of the $t+1$ numbers $|f(\epsilon^\nu)|$ is at least 1. Thus

$$\prod_{k=1}^{t} (1 + |\lambda_k|) \le (1+2)^t = 3^t \le 3^t \left| \prod_{k=1}^{t} (z_0 - \lambda_k) \right|. \tag{5}$$

If $t < h$, then for $k = t+1, \ldots, h$ we have $|\lambda_k| > 2$ and

$$\frac{1 + |\lambda_k|}{|z_0 - \lambda_k|} \le \frac{1 + |\lambda_k|}{|\lambda_k| - |z_0|} = \frac{|\lambda_k| + 1}{|\lambda_k| - 1} = 1 + \frac{2}{|\lambda_k| - 1} < 1 + \frac{2}{2-1} = 3,$$

so that
$$\prod_{k=t+1}^{h} (1 + |\lambda_k|) < 3^{h-t} \left| \prod_{k=t+1}^{h} (z_0 - \lambda_k) \right|.$$

Combining this with (5), we have

$$\prod_{k=1}^{h} (1 + |\lambda_k|) < 3^h \left| \prod_{k=1}^{h} (z_0 - \lambda_k) \right| \le 3^h ||L|| (|z_0|^h + \cdots + 1)$$

$$= 3^h (h+1) ||L|| \le 6^h ||L||.$$

THEOREM 4-3. *Suppose that $f(z)$ and $g(z)$ are polynomials with complex coefficients, of degrees n and m respectively. Suppose further that the coefficient of z^m in $g(z)$ has absolute value at least 1. Then*

$$\|f\| \leq 6^{n+m}\|fg\|.$$

Proof: Let

$$f(z) = a_0(z - \lambda_1) \cdots (z - \lambda_n),$$

$$g(z) = b_0(z - \lambda_{n+1}) \cdots (z - \lambda_{n+m}).$$

Then

$$\|f\| \leq \left\| a_0 \prod_{k=1}^{n} (z + |\lambda_k|) \right\| \leq |a_0 b_0| \cdot \prod_{k=1}^{n} (1 + |\lambda_k|)$$

$$\leq |a_0 b_0| \prod_{k=1}^{n+m} (1 + |\lambda_k|),$$

and the desired result follows from Theorem 4-2.

THEOREM 4-4. *If $f(z)$ is an arbitrary polynomial of degree n, with real coefficients, then*

$$\|f\|^m \leq (mn + 1)\|f^m\|. \tag{6}$$

Proof: Let $f(z) = a_0 + a_1 z + \cdots + a_n z^n$, and let $\|f\| = a$. The theorem is certainly true if either $|a_0| = a$ or $|a_n| = a$, since the first and last coefficients in $f^m(z)$ are the mth powers of a_0 and a_n, respectively, so that in this case $\|f^m\| \geq \|f\|^m$. If we put

$$f^*(z) = z^n f\left(\frac{1}{z}\right),$$

then clearly

$$\|f^*\| = \|f\| \quad \text{and} \quad \|(f^*)^m\| = \|(f^m)^*\|,$$

so that we can suppose, with no loss in generality, that the numerically largest of all the coefficients in $f(z)$ is a_t, where $\frac{1}{2}n \leq t < n$.

Put

$$g(z, \theta) = f(z) - a e^{i\theta} z^n,$$

and let $\alpha = \alpha(\theta)$ be the numerically largest of the zeros of $g(z, \theta)$ for each θ. The inequality (6) holds if, for some θ, $|\alpha(\theta)| \geq 1$. For

$$|f^m(\alpha)| = |a e^{i\theta} \alpha^n|^m = a^m |\alpha|^{mn},$$

while for $|\alpha| \geq 1$,

$$|f^m(\alpha)| \leq \|f^m\|(1 + |\alpha| + \cdots + |\alpha|^{mn}) \leq \|f^m\|(mn + 1)|\alpha|^{mn},$$

so that

$$a^m = ||f||^m \le ||f^m||(mn + 1).$$

We know that $|a_n| < a$. Hence if $f(1) > a$, then

$$g(1, 0) > a - a = 0, \qquad g(\infty, 0) = -\infty,$$

and $1 < |\alpha(0)| < \infty$. Similarly, if $f(1) < -a$, then

$$g(1, \pi) < -a + a = 0, \qquad g(\infty, \pi) = \infty,$$

and $1 < |\alpha(\pi)| < \infty$. This proves the theorem unless $|f(1)| \le a$, which we henceforth assume.

Now put $z = e^{i\varphi}$, so that

$$g(e^{i\varphi}, \theta) = f(e^{i\varphi}) - ae^{i(\theta + n\varphi)}.$$

If we find a φ_0 such that $|f(e^{i\varphi_0})| = a$, then θ_0 can be determined so that $g(e^{i\varphi_0}, \theta_0) = 0$; this gives $|\alpha(\theta_0)| \ge 1$ and proves the theorem. Since $|f(e^{i\varphi})|$ is a continuous function of φ, and since $|f(1)| \le a$, it suffices to prove the existence of a φ_0 such that $|f(e^{i\varphi_0})| \ge a$.

Let ϵ be a primitive $(t + 1)$th root of unity, where $|a_t| = a$ and $\frac{1}{2}n \le t < n$. Then

$$\sum_{\nu=0}^{t} \epsilon^\nu f(\epsilon^\nu) = \sum_{\nu=0}^{t} \epsilon^\nu \sum_{k=0}^{n} a_k \epsilon^{\nu k} = \sum_{k=0}^{n} a_k \sum_{\nu=0}^{t} \epsilon^{\nu(k+1)}.$$

Since $k \le n$ and $t \ge \frac{1}{2}n$, we have that $(t + 1)|(k + 1)$ if and only if $k = t$. Hence, by (4),

$$\sum_{\nu=0}^{t} \epsilon^\nu f(\epsilon^\nu) = a_t(t + 1),$$

so that for some ν,

$$|\epsilon^\nu f(\epsilon^\nu)| = |f(\epsilon^\nu)| \ge |a_t| = a.$$

The proof is complete.

THEOREM 4-5. *If $f_1(z), \ldots, f_t(z)$ are polynomials with algebraic coefficients, then*

$$\left| \prod_{\nu=1}^{t} f_\nu \right| \le \prod_{\nu=1}^{t} (1 + \deg f_\nu) \prod_{\nu=1}^{t} \overline{|f_\nu|}.$$

Proof: There is no loss in generality in supposing that

$$\deg f_1 \ge \deg f_2 \ge \cdots \ge \deg f_t.$$

The product $f_1 f_2$ is a polynomial each of whose coefficients is a sum of products of a coefficient of f_1 and a coefficient of f_2, the number of summands being at most $1 + \deg f_2$. Hence

$$\lceil f_1 f_2 \rceil = (1 + \deg f_2) \lceil f_1 \rceil \lceil f_2 \rceil.$$

Similarly,

$$\lceil f_1 f_2 f_3 \rceil \le (1 + \deg f_3) \lceil f_1 f_2 \rceil \lceil f_3 \rceil \le (1 + \deg f_3)(1 + \deg f_2) \lceil f_1 \rceil \lceil f_2 \rceil \lceil f_3 \rceil,$$

and so on.

THEOREM 4–6. *Let p and r be positive integers, with $1 \le r < p$. Suppose that $F(z_1, \ldots, z_p)$, $G(z_1, \ldots, z_r)$, and $H(z_{r+1}, \ldots, z_p)$ are polynomials with coefficients in an algebraic number field K, those of F being integers, and suppose that*

$$F(z_1, \ldots, z_p) = G(z_1, \ldots, z_r) H(z_{r+1}, \ldots, z_p).$$

Then if γ is any coefficient in F, there is a factorization $\gamma = \alpha\beta$ in K such that the coefficients in αH and βG are integers in K.

Proof: Let the coefficients in G be $\alpha_1, \ldots, \alpha_s$, and those in H be $\beta_1, \ldots, \beta_t$, in some order. Then, since the variables in G and H are disjoint, the coefficients in F are simply the products $\alpha_i \beta_j$. Since the coefficients in F are integers, all the products $\alpha_i \beta_1, \ldots, \alpha_i \beta_t$ are integers, as are all the products $\beta_j \alpha_1, \ldots, \beta_j \alpha_s$. But these two sets of numbers are just the coefficients in $\alpha_1 H$ and $\beta_1 G$.

4–3 Generalized Wronskians. Polynomials $f_0(z_1, \ldots, z_p), \ldots,$ $f_{l-1}(z_1, \ldots, z_p)$ with coefficients in an algebraic number field K are said to be *linearly dependent* if some linear combination of them, with constant coefficients in K which are not all zero, vanishes identically, and are otherwise said to be *independent*. In the case of a single independent variable, it is well known that the question of independence of a set of functions can sometimes be settled by reference to their Wronskian. For our purposes it is convenient to define this as the determinant

$$W(z) = \det\left(\frac{1}{\mu!}\frac{d^\mu}{dz^\mu} f_\nu(z)\right), \qquad \mu, \nu = 0, 1, \ldots, l-1,$$

which differs from the usual definition only in the presence of the nonzero constant factor

$$\frac{1}{0!\,1!\cdots(l-1)!}.$$

The exact relation of the behavior of the Wronskian to independence, as applied to polynomials, is indicated in the first part of the next theorem.

For functions of several variables, the situation is not quite so simple, since there are then several partial derivatives to consider. We proceed as follows. Let $\Delta_0, \Delta_1, \ldots, \Delta_\mu, \ldots, \Delta_{l-1}$ be differential operators of the form

$$\frac{1}{j_1! \cdots j_p!} \left(\frac{\partial}{\partial z_1}\right)^{j_1} \cdots \left(\frac{\partial}{\partial z_p}\right)^{j_p},$$

such that the *order* $j_1 + \cdots + j_p$ of Δ_μ does not exceed μ, for $0 \leq \mu \leq l - 1$. Then the function

$$G(z_1, \ldots, z_p) = \begin{vmatrix} \Delta_0 f_0 & \Delta_0 f_1 & \cdots & \Delta_0 f_{l-1} \\ \Delta_1 f_0 & \Delta_1 f_1 & \cdots & \Delta_1 f_{l-1} \\ \vdots & \vdots & & \vdots \\ \Delta_{l-1} f_0 & \Delta_{l-1} f_1 & \cdots & \Delta_{l-1} f_{l-1} \end{vmatrix}$$

is called a *generalized Wronskian* of $f_0, \ldots, f_{l-1}$. Except in the trivial case $p = l = 1$, there are several Δ_μ's for each μ, and hence more than one generalized Wronskian. In the case of functions of one variable, the ordinary Wronskian is that generalized Wronskian for which the order of Δ_μ is exactly μ, for $0 \leq \mu \leq l - 1$.

THEOREM 4-7. (a) *If $f_0, \ldots, f_{l-1}$ are l polynomials over K in the single variable z, whose Wronskian $W(z)$ vanishes identically, then they are dependent over K.*

(b) *If $f_0, \ldots, f_{l-1}$ are l polynomials over K in the variables $z_1, \ldots, z_p$, for which every generalized Wronskian $G_l(z_1, \ldots, z_p)$ vanishes identically, then they are dependent over K.*

Proof: (a) The proof in this case is by induction. If $l = 1$, then $W(z) = f_0(z)$, and the truth of the theorem is obvious.

Take $l > 1$, and suppose that the theorem is true for every set of $l - 1$ polynomials, $f_0, f_1, \ldots, f_{l-2}$, over K; suppose also that the Wronskian W_l of $f_0, \ldots, f_{l-1}$ vanishes identically. If $f_0, \ldots, f_{l-2}$ are dependent, so are $f_0, \ldots, f_{l-1}$, and the assertion is proved. Suppose then that $f_0, \ldots, f_{l-2}$ are independent, so that their Wronskian W_{l-1} is not identically zero. Now W_{l-1}, being a polynomial, has only finitely many zeros; let I be an interval in which it does not vanish, and take z in I. For such z, the system of equations

$$\sum_{k=0}^{l-2} f_k^{(j)}(z)y_k = f_{l-1}^{(j)}(z), \qquad j = 0, 1, \ldots, l-2, \qquad (7)$$

can be solved for the y's as rational functions of z. But then, by subtracting appropriate multiples of each column of W_l from its last column, we obtain

$$0 = 1! \cdots (l-1)! \, W_l$$

$$= \begin{vmatrix} f_0(z) & f_1(z) & \cdots & 0 \\ f_0'(z) & f_1'(z) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ f_0^{(l-1)}(z) & f_1^{(l-1)}(z) & \cdots & f_{l-1}^{(l-1)}(z) - \sum_{k=0}^{l-2} f_k^{(l-1)}(z)y_k \end{vmatrix}$$

$$= 1! \cdots (l-1)! \left(f_{l-1}^{(l-1)}(z) - \sum_{k=0}^{l-2} f_k^{(l-1)}(z)y_k \right) W_{l-1},$$

so that also

$$\sum_{k=0}^{l-2} f_k^{(l-1)}(z)y_k = f_{l-1}^{(l-1)}(z). \qquad (8)$$

Differentiating (7) gives

$$\sum_{k=0}^{l-2} f_k^{(j+1)}(z)y_k + \sum_{k=0}^{l-2} f_k^{(j)}(z)y_k' = f_{l-1}^{(j+1)}(z), \qquad j = 0, \ldots, l-2,$$

and comparison of this for $j = l-2$ with (8), and for $j = 0, \ldots, l-3$ with (7), shows that

$$\sum_{k=0}^{l-2} f_k^{(j)}(z)y_k' = 0, \qquad j = 0, \ldots, l-2.$$

Since $W_{l-1} \neq 0$, it must be that

$$y_0' = \cdots = y'_{l-2} = 0,$$

so that the y's are constants, say $y_k = c_k$, and they are clearly in K. But then the polynomial

$$\sum_{k=0}^{l-2} c_k f_k(z) - f_{l-1}(z)$$

vanishes throughout I, and therefore identically, so that the l polynomials $f_0, f_1, \ldots, f_{l-1}$ are dependent.

(b) This case is proved by contradiction. Suppose that the l polynomials $f_0(z_1, \ldots, z_p), \ldots, f_{l-1}(z_1, \ldots, z_p)$ are independent,

and suppose further that for each ν, f_ν is of degree less than k in each of its arguments, so that we can write

$$f_\nu(z_1, \ldots, z_p) = \sum_{k_1=0}^{k-1} \cdots \sum_{k_p=0}^{k-1} b_\nu(k_1, \ldots, k_p) z_1^{k_1} \cdots z_p^{k_p},$$

$$0 \leq \nu \leq l - 1.$$

Then the polynomials $f_\nu(t, t^k, t^{k^2}, \ldots, t^{k^{p-1}})$ are linearly independent. For otherwise there would be an identity in t of the form

$$\sum_{\nu=0}^{l-1} c_\nu \sum_{k_1=0}^{k-1} \cdots \sum_{k_p=0}^{k-1} b_\nu(k_1, \ldots, k_p) t^{k_1+k_2 k+\cdots+k_p k^{p-1}} = 0,$$

or

$$\sum_{k_1=0}^{k-1} \cdots \sum_{k_p=0}^{k-1} \left(\sum_{\nu=0}^{l-1} c_\nu b_\nu(k_1, \ldots, k_p) \right) t^{k_1+k_2 k+\cdots+k_p k^{p-1}} = 0,$$

and it would follow from the uniqueness of the representation of an integer to the base k that for each set of exponents $k_1, \ldots, k_p$,

$$\sum_{\nu=0}^{l-1} c_\nu b_\nu(k_1, \ldots, k_p) = 0,$$

whence

$$\sum_{\nu=0}^{l-1} c_\nu f_\nu(z_1, \ldots, z_p) = 0,$$

contrary to assumption.

We know therefore that the Wronskian

$$W(t) = \det \left(\frac{1}{\mu!} \left(\frac{d}{dt} \right)^\mu f_\nu(t, t^k, \ldots, t^{k^{p-1}}) \right), \qquad \mu, \nu = 0, \ldots, l - 1,$$

does not vanish identically. By a standard differentiation formula,

$$\frac{d}{dt} f_\nu(t, \ldots, t^{k^{p-1}}) = \sum_{j=1}^{p} \frac{\partial}{\partial z_j} f_\nu(z_1, \ldots, z_p) \bigg|_{(t, \ldots, t^{k^{p-1}})} \frac{dt^{k^{j-1}}}{dt}$$

and it follows easily by induction on μ that an operator identity

$$\left(\frac{d}{dt} \right)^\mu = \varphi_1(t) \Delta^{(1)} + \cdots + \varphi_r(t) \Delta^{(r)}$$

holds, where $\Delta^{(1)}, \ldots, \Delta^{(r)}$ are differential operators of orders not exceeding r depends only on μ and p, and $\varphi_1, \ldots, \varphi_r$ are poly- nomials with rational coefficients. Using this in the above expression

for $W(t)$, and writing the resulting determinant as a sum of other determinants, an expression for $W(t)$ of the form

$$W(t) = \psi_1(t)G_1(t, \ldots, t^{k^{p-1}}) + \cdots + \psi_s(t)G_s(t, \ldots, t^{k^{p-1}})$$

results, in which $\psi_1, \ldots, \psi_s$ are polynomials and $G_1, \ldots, G_s$ are generalized Wronskians of $f_1, \ldots, f_{l-1}$. Since $W(t)$ does not vanish identically, there is an i for which $G_i(t, \ldots, t^{k^{p-1}})$ is not identically zero, and *a fortiori* $G_i(z_1, \ldots, z_p)$ is not identically zero.

THEOREM 4–8. *Let $R(z_1, \ldots, z_p)$ be a polynomial in $p \geq 2$ variables, with integral coefficients in K such that*

$$0 < \overline{|R|} \leq B.$$

Let R be of degree at most r_j in z_j, for $j = 1, \ldots, p$. Then there is an l in Z with

$$1 \leq l \leq r_p + 1, \tag{9}$$

there is an integer β in K, and there are differential operators $\Delta_0, \ldots, \Delta_{l-1}$ on the variables $z_1, \ldots, z_{p-1}$, of orders at most $0, \ldots, l-1$, respectively, such that if

$$F(z_1, \ldots, z_p) = \beta \det\left(\Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial z_p}\right)^\nu R\right), \quad \mu, \nu = 0, \ldots, l-1, \tag{10}$$

then

    (a) *F has integral coefficients in K and is not identically zero;*
    (b) *a decomposition*

$$F(z_1, \ldots, z_p) = U(z_1, \ldots, z_{p-1})V(z_p) \tag{11}$$

holds, where U and V have integral coefficients in K, U is of degree at most lr_j in z_j for $j = 1, \ldots, p-1$, and V is of degree at most lr_p in z_p;

    (c) *the following bound holds:*

$$\overline{|F|} \leq \{(r_1 + 1) \cdots (r_p + 1)\}^{2l} 2^{2(r_1 + \cdots + r_p)l} l!^2 B^{2l}.$$

Proof: Write R as a polynomial in z_p:

$$R(z_1, \ldots, z_p) = \sum_{\varkappa=0}^{r_p} S_\varkappa(z_1, \ldots, z_{p-1})z_p^\varkappa.$$

The polynomials $S_\varkappa$ need not be independent; let $\psi_\nu(z_1, \ldots, z_{p-1})$, for $\nu = 0, \ldots, l-1$, be a maximal set of independent polynomials

among the $S_\varkappa$, so that $1 \le l \le r_p + 1$. Then there are constants $\beta_{\nu\varkappa}$ in K such that for $\varkappa = 0, \ldots, r_p$,

$$S_\varkappa(z_1, \ldots, z_{p-1}) = \sum_{\nu=0}^{l-1} \beta_{\nu\varkappa}\psi_\nu(z_1, \ldots, z_{p-1}). \qquad (12)$$

If we put

$$\varphi_\nu(z_p) = \sum_{\varkappa=0}^{r_p} \beta_{\nu\varkappa}z_p^\varkappa, \qquad \nu = 0, \ldots, l-1,$$

then

$$R(z_1, \ldots, z_p) = \sum_{\nu=0}^{l-1} \psi_\nu(z_1, \ldots, z_{p-1})\, \varphi_\nu(z_p), \qquad (13)$$

and $\varphi_0, \ldots, \varphi_{l-1}$ are independent. For if $\delta_0, \ldots, \delta_{l-1}$ are constants such that

$$\delta_0\varphi_0(z_p) + \cdots + \delta_{l-1}\varphi_{l-1}(z_p) = 0,$$

the coefficient of each power of z_p must be zero, so that

$$\delta_0\beta_{0\varkappa} + \cdots + \delta_{l-1}\beta_{l-1,\varkappa} = 0 \qquad (14)$$

for $\varkappa = 0, \ldots, r_p$. For fixed ν_0 with $0 \le \nu_0 \le l - 1$, choose $\varkappa_0$ so that $S_{\varkappa_0}(z_1, \ldots, z_{p-1}) = \psi_{\nu_0}(z_1, \ldots, z_{p-1})$; this is possible since the ψ's are a subset of the S's. Then (12) shows that

$$\beta_{\nu\varkappa_0} = \begin{cases} 1 & \text{if } \nu = \nu_0, \\ 0 & \text{if } \nu \ne \nu_0. \end{cases}$$

Choosing $\varkappa = \varkappa_0$ in (14), we obtain $\delta_{\nu_0} = 0$. Since ν_0 is arbitrary, every $\delta_i = 0$.

Let $W(z_p)$ be the Wronskian of $\varphi_0, \ldots, \varphi_{l-1}$; it is a polynomial with coefficients in K, and it does not vanish identically. Let $G(z_1, \ldots, z_{p-1})$ be some generalized Wronskian of $\psi_0, \ldots, \psi_{l-1}$ which is not identically zero. Then

$$W(z_p) = \det\left(\frac{1}{\mu!}\left(\frac{d}{dz_p}\right)^\mu \varphi_\nu(z_p)\right), \qquad \mu, \nu = 0, \ldots, l-1,$$

$$G(z_1, \ldots, z_{p-1}) = \det\left(\Delta_\mu\psi_\nu(z_1, \ldots, z_{p-1})\right),$$

where $\Delta_0, \ldots, \Delta_{l-1}$ are differential operators on $z_1, \ldots, z_{p-1}$, of orders at most $0, \ldots, l - 1$ respectively. Taking the row-by-row product of G and W, we obtain

$$GW = \det\left(\sum_{\rho=0}^{l-1} \Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial z_p}\right)^\nu \varphi_\rho(z_p)\, \psi_\rho(z_1, \ldots, z_{p-1})\right),$$

or
$$GW = \det\left(\Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial z_p}\right)^\nu R\right). \tag{15}$$

Since W is a determinant of order l whose elements are polynomials in z_p of degrees at most r_p, it is clear that $\deg W \le lr_p$. Similarly, G is of degree at most lr_j in z_j, for $j = 1, \ldots, p-1$.

In the expression (15) for GW, we can write R as the sum of $(r_1 + 1) \cdots (r_p + 1)$ terms of the form

$$\alpha_{s_1 \cdots s_p} z_1^{s_1} \cdots z_p^{s_p}.$$

The determinant can then be written as a sum of

$$((r_1 + 1) \cdots (r_p + 1))^l$$

new determinants, each having entries of the form

$$\alpha_{s_1 \cdots s_p} \Delta_\mu \frac{1}{\nu!}\left(\frac{\partial}{\partial z_p}\right)^\nu z_1^{s_1} \cdots z_p^{s_p} = a\alpha_{s_1 \cdots s_p} z_1^{t_1} \cdots z_p^{t_p},$$

in which $t_j \le s_j$ for $j = 1, \ldots, p$. Here

$$a = \binom{s_1}{t_1} \cdots \binom{s_p}{t_p} \le 2^{s_1 + \cdots + s_p} \le 2^{r_1 + \cdots + r_p}.$$

Thus the entries of each new determinant are such that the maxima of the absolute values of their conjugates do not exceed

$$2^{r_1 + \cdots + r_p} B,$$

and hence

$$\overline{|GW|} \le ((r_1 + 1) \cdots (r_p + 1))^l l! 2^{(r_1 + \cdots + r_p)l} B^l.$$

The coefficients in GW are integers in K. It follows from Theorem 4-6 that if β is any one of them which is not zero, there is a factorization $\beta = \beta_1\beta_2$ in K such that $\beta_1 G = U$ and $\beta_2 W = V$ have integral coefficients in K, and

$$\beta GW = F = UV.$$

By the bound just obtained for $\overline{|GW|}$, we have

$$0 < \overline{|F|} \le \overline{|GW|}^2 \le ((r_1 + 1) \cdots (r_p + 1))^{2l} l!^2 2^{2(r_1 + \cdots + r_p)l} B^{2l}.$$

4-4 The index. Let $P(z_1, \ldots, z_p)$ be any polynomial in p variables which does not vanish identically. Let $\alpha_1, \ldots, \alpha_p$ be any complex numbers, and let $r_1, \ldots, r_p$ be any positive numbers. We

define the *index θ of P at the point* $(\alpha_1, \ldots, \alpha_p)$ *relative to* $r_1, \ldots, r_p$ as follows. Expand $P(\alpha_1 + y_1, \ldots, \alpha_p + y_p)$ as a polynomial in $y_1, \ldots, y_p$, say

$$P(\alpha_1 + y_1, \ldots, \alpha_p + y_p) = \sum_{j_1=0}^{\infty} \cdots \sum_{j_p=0}^{\infty} c(j_1, \ldots, j_p) y_1^{j_1} \cdots y_p^{j_p}.$$

Then

$$\theta = \min \left(\frac{j_1}{r_1} + \cdots + \frac{j_p}{r_p} \right),$$

the minimum being extended over all sets of non-negative integers $j_1, \ldots, j_p$ for which $c(j_1, \ldots, j_p) \neq 0$, or, equivalently, for which

$$\left(\frac{\partial}{\partial z_1} \right)^{j_1} \cdots \left(\frac{\partial}{\partial z_p} \right)^{j_p} P(\alpha_1, \ldots, \alpha_p) \neq 0.$$

Note that $\theta \geq 0$ always, and that $\theta = 0$ if and only if

$$P(\alpha_1, \ldots, \alpha_p) \neq 0.$$

Moreover, if any derived polynomial

$$\left(\frac{\partial}{\partial z_1} \right)^{k_1} \cdots \left(\frac{\partial}{\partial z_p} \right)^{k_p} P(z_1, \ldots, z_p)$$

is not identically zero, it is clear that its index at $(\alpha_1, \ldots, \alpha_p)$ relative to $r_1, \ldots, r_p$ is at least

$$\theta - \frac{k_1}{r_1} - \cdots - \frac{k_p}{r_p}.$$

The following properties, which we list in a theorem for later reference, are also immediate consequences of the definition.

THEOREM 4–9. *Let* $P(z_1, \ldots, z_p)$ *and* $Q(z_1, \ldots, z_p)$ *be polynomials, neither of which vanishes identically. Then if we consider indices formed at the same point* $(\alpha_1, \ldots, \alpha_p)$ *relative to the same numbers* $r_1, \ldots, r_p$, *the following relations hold:*

$$\text{index } (P + Q) \geq \min (\text{index } P, \text{index } Q), \tag{16}$$

$$\text{index } PQ = \text{index } P + \text{index } Q. \tag{17}$$

Equation (17) *remains true if P is a polynomial in $z_1, \ldots, z_{p-1}$ only, and Q is a polynomial in z_p only, provided that the index of P is taken at $(\alpha_1, \ldots, \alpha_{p-1})$ relative to $r_1, \ldots, r_{p-1}$, and that of Q at α_p relative to r_p.*

Now let $r_1, \ldots, r_m$ be positive integers, and suppose that $B \geq 1$. We consider the set $\mathcal{R}_m = \mathcal{R}_m(B; r_1, \ldots, r_m)$ of polynomials $R(z_1, \ldots, z_m)$ which satisfy the following conditions:

(a) R has integral coefficients in K, and is not identically zero.

(b) R is of degree at most r_j in z_j, for $j = 1, \ldots, m$.

(c) $\lceil R \rceil \leq B$.

Let $\zeta_1, \ldots, \zeta_m$ be algebraic numbers (not necessarily in K) of heights $H(\zeta_1) = q_1, \ldots, H(\zeta_m) = q_m$. Let $\theta(R)$ denote the index of $R(z_1, \ldots, z_m)$ at the point $(\zeta_1, \ldots, \zeta_m)$ relative to $r_1, \ldots, r_m$. Our object in the present section is to obtain, under certain conditions, an upper bound for $\theta(R)$ in terms of $B, q_1, \ldots, q_m, r_1, \ldots, r_m$. We therefore define

$$\Theta_m(B; q_1, \ldots, q_m; r_1, \ldots, r_m) = \sup \theta(R), \tag{18}$$

the supremum, or least upper bound, being taken over all R in $\mathcal{R}$ and all integers $\zeta_1, \ldots, \zeta_m$ of heights $q_1, \ldots, q_m$, respectively.

The double significance of $r_1, \ldots, r_m$ in the definition (18) should be noted; these numbers occur both in the definition of the index and in condition (b) above.

We proceed by induction on m. In Theorem 4–10 the case $m = 1$ is treated, in Theorem 4–11 there is given a recurrence relation between Θ_{m-1} and Θ_m, and in Theorem 4–12 an explicit bound is obtained.

THEOREM 4–10.

$$\Theta_1(B; q_1; r_1) \leq \frac{3N(N+1)}{\log q_1} + \frac{N \log B}{r_1 \log q_1}.$$

Proof: Let the defining polynomial of ζ_1 be

$$\chi(z_1) = d_0 z_1^h + \cdots + d_h, \qquad d_0 \neq 0,$$

where $d_0, \ldots, d_h$ are relatively prime rational integers, so that

$$\|\chi\| = H(\zeta_1) = q_1 = \max(|d_0|, \ldots, |d_h|).$$

Each polynomial R in $\mathcal{R}_1$ has integral coefficients in K; regarding these coefficients as polynomials in a single primitive element, we can obtain other polynomials from R by successively replacing this primitive element throughout by its various conjugates. Let R^* be the product of these N polynomials. By the Symmetric Function

Theorem, R^* has coefficients in Z. Also, $\deg R^* = Nr_1$, and by Theorem 4–5,

$$\|R^*\| \le (1 + r_1)^N B^N.$$

By the definition of the index, $R(z_1)$ is divisible by $(z_1 - \zeta_1)^{r_1\theta}$, and the same is therefore true of $R^*(z_1)$. Since $R^*(z_1)$ has coefficients in Z, it is divisible by $\chi^{r_1\theta}$. One consequence of this fact is that $hr_1\theta \le Nr_1$. Also, it follows from Theorem 4–3 that

$$\|\chi^{r_1\theta}\| \le 6^{Nr_1}\|R^*\|,$$

and, by Theorem 4–4,

$$
\begin{aligned}
q_1{}^{r_1\theta} = \|\chi\|^{r_1\theta} &\le (hr_1\theta + 1)6^{Nr_1}\|R^*\| \\
&\le (Nr_1 + 1)^{N+1}6^{Nr_1}B^N < 2^{Nr_1(N+1)}6^{Nr_1}B^N \\
&< 12^{N(N+1)r_1}B^N.
\end{aligned}
$$

Hence

$$\theta < \frac{N(N+1)\log 12}{\log q_1} + \frac{N \log B}{r_1 \log q_1},$$

and the theorem follows from the fact that $\log 12 < 3$.

THEOREM 4–11. *Let $p \ge 2$ be a positive integer, let $r_1, \ldots, r_p$ be positive integers such that*

$$r_p > 10\delta^{-1}, \qquad \frac{r_{j-1}}{r_j} > \delta^{-1}, \qquad \text{for } j = 2, \ldots, p, \qquad (19)$$

where $0 < \delta < 1$, and let $q_1, \ldots, q_p$ be positive integers. Then

$$\Theta_p(B; q_1, \ldots, q_p; r_1, \ldots, r_p) \le 2 \max (\Phi + \Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}}), \qquad (20)$$

where the maximum is taken over integers l satisfying

$$1 \le l \le r_p + 1, \qquad (21)$$

and where

$$\Phi = \Theta_1(M; q_p; lr_p) + \Theta_{p-1}(M; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}) \qquad (22)$$

and

$$M = (r_1 + 1)^{2pl}2^{2r_1pl}l!^2 B^{2l}. \qquad (23)$$

Proof: Let $R(z_1, \ldots, z_p)$ be any polynomial of the class $\mathcal{R}_p(B; r_1, \ldots, r_p)$ and let $\zeta_1, \ldots, \zeta_p$ be algebraic numbers of heights $q_1, \ldots, q_p$ respectively. Then R satisfies the hypotheses of Theorem 4–8, so that there are numbers l and β and a polynomial $F(z_1, \ldots, z_p)$

having the properties listed there. By Theorem 4–8,

$$\lceil F \rceil \le ((r_1 + 1) \cdots (r_p + 1))^{2l} 2^{2(r_1 + \cdots + r_p)l} l!^2 B^{2l}$$

and hence

$$\lceil F \rceil < (r_1 + 1)^{2pl} 2^{2r_1 pl} l!^2 B^{2l} = M,$$

since $r_1 > r_2 > \cdots > r_p$ by (19). From the factorization

$$F(z_1, \ldots, z_p) = U(z_1, \ldots, z_{p-1}) V(z_p)$$

and the fact that the arguments of U and V are disjoint, it follows that also

$$\lceil U \rceil < M, \qquad \lceil V \rceil < M.$$

The polynomial $U(z_1, \ldots, z_{p-1})$ has degree at most lr_j in z_j, for $j = 1, \ldots, p - 1$. It is therefore an element of the class

$$\mathcal{R}_{p-1}(M; lr_1, \ldots, lr_{p-1}).$$

Hence, its index at $(\zeta_1, \ldots, \zeta_{p-1})$ relative to $lr_1, \ldots, lr_{p-1}$ is at most

$$\Theta_{p-1}(M; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}).$$

It follows from the definition of the index that the index of U at that point relative to $r_1, \ldots, r_{p-1}$ is at most

$$l\Theta_{p-1}(M; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}).$$

Similarly, $V(z_p)$ is an element of the class $\mathcal{R}_1(M; lr_p)$, and its index at ζ_p relative to r_p is at most

$$l\Theta_1(M; q_p; lr_p).$$

By the last sentence of Theorem 4–9, the index of $F = UV$ at $(\zeta_1, \ldots, \zeta_p)$ relative to $r_1, \ldots, r_p$ is the sum of the indices of U and V, so that

$$\text{index } F \le l\Phi, \tag{24}$$

where Φ is defined in (22).

We now deduce from the determinantal representation of F in equation (10) a lower bound for the index of F in terms of the index θ of R. Consider first any differential operator of the form

$$\Delta = \frac{1}{i_1! \cdots i_{p-1}!} \left(\frac{\partial}{\partial z_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial z_{p-1}}\right)^{i_{p-1}},$$

of order

$$w = i_1 + \cdots + i_{p-1} \le l - 1.$$

If the polynomial

$$\Delta \frac{1}{\nu!} \left(\frac{\partial}{\partial z_p}\right)^\nu R(z_1, \ldots, z_p)$$

does not vanish identically, its index at $(\zeta_1, \ldots, \zeta_p)$ relative to $r_1, \ldots, r_p$ is at least

$$\theta - \frac{i_1}{r_1} - \cdots - \frac{i_{p-1}}{r_{p-1}} - \frac{\nu}{r_p} \geq \theta - \frac{w}{r_{p-1}} - \frac{\nu}{r_p}.$$

Now

$$\frac{w}{r_{p-1}} \leq \frac{l-1}{r_{p-1}} \leq \frac{r_p}{r_{p-1}} < \delta,$$

by the inequalities (21) and (19). Hence, since the index is non-negative, it must be at least

$$\max\left(0, \theta - \delta - \frac{\nu}{r_p}\right) \geq \max\left(0, \theta - \frac{\nu}{r_p}\right) - \delta.$$

If we expand the determinant on the right side of (10), we obtain for F a sum of $l!$ terms, a typical term being

$$\pm \beta(\Delta_{\mu_0}R)\left(\Delta_{\mu_1}\frac{1}{1!}\frac{\partial}{\partial z_p}R\right)\cdots\left(\Delta_{\mu_{l-1}}\frac{1}{(l-1)!}\left(\frac{\partial}{\partial z_p}\right)^{l-1}R\right),$$

where $\Delta_{\mu_0}, \ldots, \Delta_{\mu_{l-1}}$ are differential operators on $z_1, \ldots, z_{p-1}$ whose orders are at most $l-1$. By Theorem 4–9, the index of such a term, if it does not vanish identically, is at least

$$\sum_{\nu=0}^{l-1} \max\left(0, \theta - \frac{\nu}{r_p}\right) - l\delta.$$

Since F is a sum of such terms, it follows from Theorem 4–9 again that

$$\text{index } F \geq \sum_{\nu=0}^{l-1} \max\left(0, \theta - \frac{\nu}{r_p}\right) - l\delta.$$

We may suppose that $\theta r_p > 10$, since otherwise

$$\theta < 10 r_p^{-1} < \delta < 2\delta^{\frac{1}{2}}$$

and the desired inequality for θ then holds. Under this supposition, $[\theta r_p]^2 > 2\theta^2 r_p{}^2/3$. Hence if $\theta r_p < l$, we have

$$\sum_{\nu=0}^{l-1} \max\left(0, \theta - \frac{\nu}{r_p}\right) = r_p^{-1} \sum_{\nu=0}^{[\theta r_p]} (\theta r_p - \nu)$$

$$\geq \tfrac{1}{2} r_p^{-1} [\theta r_p]^2$$

$$\geq \tfrac{1}{3} \theta^2 r_p,$$

while if $\theta r_p \geq l$, then

$$\sum_{\nu=0}^{l-1} \max\left(0, \theta - \frac{\nu}{r_p}\right) = \sum_{\nu=0}^{l-1}\left(\theta - \frac{\nu}{r_p}\right) \geq \frac{1}{2} l\theta.$$

Hence

$$\text{index } F \geq \min\left(\tfrac{1}{2}l\theta, \tfrac{1}{3}r_p\theta^2\right) - l\delta. \tag{25}$$

Combining (24) and (25), we obtain

$$\min\left(\tfrac{1}{2}l\theta, \tfrac{1}{3}r_p\theta^2\right) \leq l(\Phi + \delta).$$

Thus either $\theta < 2(\Phi + \delta)$, in which case θ satisfies the desired inequality, or

$$\tfrac{1}{3}r_p\theta^2 \leq l(\Phi + \delta) \leq (r_p + 1)(\Phi + \delta).$$

Since $r_p + 1 < 4r_p/3$ by (19), this gives

$$\theta \leq 2(\Phi + \delta)^{\frac{1}{2}} \leq 2(\Phi^{\frac{1}{2}} + \delta^{\frac{1}{2}}),$$

and the proof is complete.

THEOREM 4–12. *Let m be a positive integer, and suppose that*

$$0 < \delta < \frac{1}{m2^m(N+1)^2}. \tag{26}$$

Let $r_1, \ldots, r_m$ be positive integers such that

$$r_m > 10\delta^{-1}, \qquad \frac{r_{j-1}}{r_j} > \delta^{-1}, \qquad \text{for } j = 2, \ldots, m. \tag{27}$$

Let $q_1, \ldots, q_m$ be positive integers such that

$$\log q_1 > 2\delta^{-1} m(2m+1), \tag{28}$$

$$r_j \log q_j \geq r_1 \log q_1, \qquad \text{for } j = 2, \ldots, m, \tag{29}$$

$$\log q_1 > 3\delta^{-1}N(N+1). \tag{30}$$

Then

$$\Theta_m(q_1^{\delta r_1}; q_1, \ldots, q_m; r_1, \ldots, r_m) < 10^m \delta^{(\frac{1}{4})^m}. \tag{31}$$

Proof: The proof is by induction on m. For $m = 1$, we apply Theorem 4–10, together with the inequalities (30) and (26), and obtain

$$\theta_1 \left(q_1^{\delta r_1}; q_1; r_1\right) < \frac{3N(N+1)}{\log q_1} + \frac{N \log \left(q_1^{\delta r_1}\right)}{r_1 \log q_1} < (N+1)\delta < 10\delta^{\frac{1}{2}},$$

which is the desired inequality.

Now suppose that $p \geq 2$ is an integer, and that the theorem holds when $m = p - 1$. When $m = p$, the hypotheses of the present theorem are more stringent than those of Theorem 4–11, so that the latter is applicable here. We must estimate M and Φ.

We have

$$M = (r_1 + 1)^{2pl} 2^{2r_1 pl} l!^2 B^{2l} \leq \left((r_1 + 1)^{2p} 2^{2r_1 p} l^2 q_1^{2\delta r_1}\right)^l.$$

Since $l \leq r_p + 1 < r_1 + 1 \leq 2^{r_1}$, it follows that

$$M < \left(2^{(4p+2)r_1} q_1^{2\delta r_1}\right)^l < \left(e^{(4p+2)r_1} q_1^{2\delta r_1}\right)^l.$$

By (28) with $m = p$, we have $4p + 2 < \delta p^{-1} \log q_1$, so that

$$M < q_1^{\delta_1 l r_1},$$

where
$$\delta_1 = 2\delta(1 + p^{-1}). \tag{32}$$

Thus
$$\theta_1(M; q_p; l r_p) \leq \theta_1(q^{\delta_1 l r_1}; q_p; l r_p) \tag{33}$$
and

$$\theta_{p-1}(M; q_1, \ldots, q_{p-1}; l r_1, \ldots, l r_{p-1})$$
$$\leq \theta_{p-1}(q_1^{\delta_1 l r_1}; q_1, \ldots, q_{p-1}; l r_1, \ldots, l r_{p-1}). \tag{34}$$

Moreover, (32), together with the inequality (26) with $m = p$, implies that

$$\delta_1 < \frac{1 + p^{-1}}{p 2^{p-1}(N+1)^2} < \frac{1}{(p-1)2^{p-1}(N+1)^2}. \tag{35}$$

In particular, $(N+1)\delta_1 < \delta_1^{\frac{1}{2}}$.

It follows from (30), and the fact that $q_p > q_1$, that

$$\log q_p > 3\delta^{-1} N(N+1).$$

Hence by Theorem 4–10, the right side of (33) does not exceed

$$\delta + \frac{N\delta_1 l r_1 \log q_1}{l r_p \log q_p} \leq \delta + N\delta_1 < (N+1)\delta_1 < \delta_1^{\frac{1}{2}};$$

here we have used (29).

To estimate the right side of (34), we use the induction hypothesis, that the theorem holds when $m = p - 1$. The conditions of the theorem are satisfied for $m = p - 1$, if we replace δ by δ_1 and $r_1, \ldots, r_{p-1}$ by $lr_1, \ldots, lr_{p-1}$; since $\delta_1 > \delta$, this is obvious for all the relations but (26), which has already been verified in (35). It follows that

$$\Theta_{p-1}(q_1{}^{\delta_1 l r_1}; q_1, \ldots, q_{p-1}; lr_1, \ldots, lr_{p-1}) < 10^{p-1}\delta_1{}^{(\frac{1}{2})^{p-1}}.$$

Hence, since $\delta_1 < 4\delta$, the two results just proved imply that

$$\Phi < 2\delta^{\frac{1}{2}} + 2(10^{p-1}\delta^{(\frac{1}{2})^{p-1}}) < 3(10^{p-1}\delta^{(\frac{1}{2})^{p-1}}).$$

Finally, (20) gives

$$\Theta_p(q_1{}^{\delta r_1}; q_1, \ldots, q_p; r_1, \ldots, r_p)$$

$$< 2\{3(10^{p-1}\delta^{(\frac{1}{2})^{p-1}}) + 3^{\frac{1}{2}}10^{\frac{1}{2}(p-1)}\delta^{(\frac{1}{2})^p} + \delta^{\frac{1}{2}}\}$$

$$< 2\left(\frac{3}{10} + \frac{3^{\frac{1}{2}}}{10^{\frac{3}{2}}} + \frac{1}{10^2}\right)10^p\delta^{(\frac{1}{2})^p} < 10^p\delta^{(\frac{1}{2})^p}.$$

4–5 A combinatorial lemma

THEOREM 4–13. *If $r_1, \ldots, r_m$ are any positive integers, and $\lambda > 0$, then the number $A_m(\lambda)$ of sets of integers $j_1, \ldots, j_m$ which satisfy the inequalities*

$$0 \leq j_1 \leq r_1, \quad \ldots, \quad 0 \leq j_m \leq r_m,$$

$$\frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \leq \frac{1}{2}(m - \lambda)$$

does not exceed

$$2m^{\frac{1}{2}}\lambda^{-1}(r_1 + 1) \cdots (r_m + 1).$$

Proof: We proceed by induction on m. The theorem holds for $m = 1$, since the number of integers j_1 such that

$$0 \leq j_1 \leq r_1, \quad j_1 \leq \tfrac{1}{2}(1 - \lambda)r_1$$

is at most $r_1 + 1$, and is 0 if $\lambda > 1$.

Now suppose $m > 1$. The result is trivial if $\lambda \leq 2m^{\frac{1}{2}}$, since then the conditions on the individual j's give an improvement of the desired upper bound. Hence we may suppose that $\lambda > 2m^{\frac{1}{2}}$. If we fix j_m, we must count the sets of integers $j_1, \ldots, j_{m-1}$ such that

$$0 \le j_1 \le r_1, \qquad \ldots, \qquad 0 \le j_{m-1} \le r_{m-1},$$

$$\frac{j_1}{r_1} + \cdots + \frac{j_{m-1}}{r_{m-1}} \le \frac{1}{2}\left(m - \lambda - \frac{2j_m}{r_m}\right).$$

Putting

$$m - \lambda - \frac{2j_m}{r_m} = (m - 1) - \lambda',$$

or, what is the same thing,

$$\lambda' = \lambda'(j_m) = \lambda - 1 + \frac{2j_m}{r_m},$$

we see that

$$A_m(\lambda) = \sum_{j_m=0}^{r_m} A_{m-1}(\lambda'(j_m)).$$

By the induction hypothesis,

$$A_m(\lambda) \le 2(m-1)^{\frac{1}{2}}(r_1 + 1) \cdots (r_{m-1} + 1) \sum_{j=0}^{r_m} \left(\lambda - 1 + \frac{2j}{r_m}\right)^{-1},$$

and it suffices to prove that

$$\sum_{j=0}^{r} \left(\lambda - 1 + \frac{2j}{r}\right)^{-1} \le \lambda^{-1}(m-1)^{-\frac{1}{2}}m^{\frac{1}{2}}(r+1)$$

for all positive integers r and m, if $\lambda > 2m^{\frac{1}{2}}$.

If r is even, we put $j = \frac{1}{2}r + k$ and obtain the sum

$$\sum_{k=-\frac{1}{2}r}^{\frac{1}{2}r} \left(\lambda + \frac{2k}{r}\right)^{-1} = \lambda^{-1} + \sum_{k=1}^{\frac{1}{2}r} \left\{\left(\lambda + \frac{2k}{r}\right)^{-1} + \left(\lambda - \frac{2k}{r}\right)^{-1}\right\}$$

$$= \lambda^{-1} + \sum_{k=1}^{\frac{1}{2}r} 2\lambda\left(\lambda^2 - \frac{4k^2}{r^2}\right)^{-1}$$

$$\le \lambda^{-1} + 2\lambda \sum_{k=1}^{\frac{1}{2}r} (\lambda^2 - 1)^{-1}$$

$$= \lambda^{-1} + 2\lambda^{-1} \sum_{k=1}^{\frac{1}{2}r} (1 - \lambda^{-2})^{-1}$$

$$\le \lambda^{-1}(r+1)(1 - \lambda^{-2})^{-1}.$$

Since $1 - \lambda^{-2} > 1 - m^{-1}/4 > (1 - m^{-1})^{\frac{1}{2}}$, we have the desired inequality.

If r is odd, we put $j = (r-1)/2 + k$ and obtain the sum

$$\sum_{k=-\frac{1}{2}(r-1)}^{\frac{1}{2}(r+1)} \left(\lambda + \frac{2k-1}{r}\right)^{-1}$$

$$= \sum_{k=1}^{\frac{1}{2}(r+1)} \left\{ \left(\lambda + \frac{2k-1}{r}\right)^{-1} + \left(\lambda - \frac{2k-1}{r}\right)^{-1} \right\}$$

$$= 2\lambda \sum_{k=1}^{\frac{1}{2}(r+1)} \left(\lambda^2 - \frac{(2k-1)^2}{r^2}\right)^{-1}$$

$$\leq \lambda(\lambda^2 - 1)^{-1}(r+1),$$

and the result is as before.

4–6 The approximation polynomial. Let α be an algebraic integer of degree $n \geq 2$ over K, so that α is a zero of a polynomial which has integral coefficients in K and which cannot be factored into a product of such polynomials of positive degrees. Let $L = K(\alpha)$ be the field obtained by adjoining α to K. Finally, let $\omega_1, \ldots, \omega_N$ be an integral basis for K, and put

$$\lceil \alpha \rceil = b_1, \qquad \max\left(\lceil \omega_1 \rceil, \ldots, \lceil \omega_N \rceil\right) = b_2. \tag{36}$$

In the remainder of the proof we shall be concerned with a single set of values of $m, \delta, q_1, \zeta_1, \ldots, q_m, \zeta_m, r_1, \ldots, r_m$, which will be chosen later in the order just specified. The choice will be made so as to satisfy the following conditions:

$$0 < \delta < m^{-1}2^{-m}(N+1)^{-2}, \tag{37}$$

$$10^m \delta^{(\frac{1}{2})^m} + 2(1 + 3\delta)nm^{\frac{1}{2}} < \frac{m}{2}, \tag{38}$$

$$r_m > 10\delta^{-1}, \qquad \frac{r_{j-1}}{r_j} > \delta^{-1}, \qquad \text{for } j = 2, \ldots, m, \tag{39}$$

$$\delta^2 \log q_1 > 2m + 1 + m \log(b_1 + 1) + 4b_2 N, \tag{40}$$

$$r_j \log q_j \geq r_1 \log q_1, \qquad \text{for } j = 2, \ldots, m, \tag{41}$$

$$\log q_1 > 3\delta^{-1} N(N+1). \tag{42}$$

Notice that these conditions imply those of Theorem 4–12, since (37) and (40) together imply that $\delta \log q_1 > 2m(2m + 1)$.

Define λ, μ, η, B_1 by the equations

$$\lambda = 4(1 + 3\delta)nm^{\frac{1}{2}}, \tag{43}$$

$$\mu = \tfrac{1}{2}(m - \lambda), \tag{44}$$

$$\eta = 10^m \delta^{(\frac{1}{2})^m}, \tag{45}$$

$$B_1 = [q_1{}^{\delta r_1}]. \tag{46}$$

Then (38) is equivalent to

$$\eta < \mu. \tag{47}$$

Also,

$$q_1{}^{\frac{1}{2}\delta r_1} < B_1,$$

since $\sqrt{x} < x - 1 < [x]$ for all $x > (3 + \sqrt{5})/2$, and

$$q^{\delta r_1} > q^{\delta^2 r_1} > e^{(2m+1)r_1} > e^{3rm} > e^{30}.$$

We come now to the main lemma, which will be the only one to which reference is made in the eventual proof of the Thue-Siegel-Roth theorem.

THEOREM 4–14. *Suppose that the conditions (37) through (42) are satisfied, and suppose that $\zeta_1, \ldots, \zeta_m$ are algebraic numbers of heights $q_1, \ldots, q_m$, respectively. Then there exists a polynomial $Q(z_1, \ldots, z_m)$ with integral coefficients in K and of degree at most r_j in z_j, for $j = 1, \ldots, m$, such that*

(a) *the index of Q at the point $(\alpha, \ldots, \alpha)$ relative to $r_1, \ldots, r_m$ is at least $\mu - \eta$;*

(b) $Q(\zeta_1, \ldots, \zeta_m) \neq 0$;

(c) *for all derivatives*

$$Q_{i_1 \cdots i_m}(z_1, \ldots, z_m) = \frac{1}{i_1! \cdots i_m!}\left(\frac{\partial}{\partial z_1}\right)^{i_1} \cdots \left(\frac{\partial}{\partial z_m}\right)^{i_m} Q,$$

where $i_1, \ldots, i_m$ are non-negative integers, the inequality

$$|Q_{i_1 \cdots i_m}(z_1, \ldots, z_m)| < B_1^{1+3\delta}(1 + |z_1|)^{r_1} \cdots (1 + |z_m|)^{r_m}$$

holds, and the corresponding inequality also holds if the coefficients in Q are replaced by their respective field conjugates.

Proof: Let $c_1, \ldots, c_N$ range independently over the non-negative rational integers not exceeding B_1, and let C be the set of integers of K of the form

$$c_1\omega_1 + \cdots + c_N\omega_N.$$

The number of elements of C is $(1 + B_1)^N$, and if we put

$$(1 + r_1) \cdots (1 + r_m) = r,$$

there are

$$(1 + B_1)^{Nr} \tag{48}$$

distinct polynomials

$$P(z_1, \ldots, z_m) = \sum_{s_1=0}^{r_1} \cdots \sum_{s_m=0}^{r_m} \gamma(s_1, \ldots, s_m) z_1^{s_1} \cdots z_m^{s_m}$$

whose coefficients $\gamma(s_1, \ldots, s_m)$ belong to C. For $\gamma(s_1, \cdots, s_m)$ in C,

$$\overline{|\gamma(s_1, \ldots, s_m)|} \leq b_2 B_1 N, \tag{49}$$

and if we put

$$P_{j_1 \cdots j_m}(z_1, \ldots, z_m)$$

$$= \frac{1}{j_1! \cdots j_m!} \left(\frac{\partial}{\partial z_1}\right)^{j_1} \cdots \left(\frac{\partial}{\partial z_m}\right)^{j_m} P(z_1, \ldots, z_m)$$

$$= \sum_{s_1=0}^{r_1} \cdots \sum_{s_m=0}^{r_m} \gamma(s_1, \ldots, s_m) \binom{s_1}{j_1} \cdots \binom{s_m}{j_m} z_1^{s_1-j_1} \cdots z_m^{s_m-j_m},$$

then

$$\overline{|P_{j_1 \cdots j_m}|} \leq 2^{r_1 + \cdots + r_m} b_2 B_1 N \leq b_2 N \, 2^{mr_1} B_1 < b_2 N B_1^{1+\delta},$$

since $mr_1 \log 2 < \frac{1}{2}\delta^2 r_1 \log q_1$ by (40). Now replace all of $z_1, \ldots, z_m$ by α. Since the total number of terms is at most r, and since, by (40),

$$r = (r_1 + 1) \cdots (r_m + 1) \leq 2^{r_1 + \cdots + r_m} \leq (b_1 + 1)^{mr_1} < B_1^{\delta}, \tag{50}$$

we obtain the bound

$$\overline{|P_{j_1 \cdots j_m}(\alpha, \ldots, \alpha)|} \leq b_2 N B_1^{1+\delta} r b_1^{r_1 + \cdots + r_m}$$
$$\leq b_2 N B_1^{1+3\delta}.$$

Let ϑ be a primitive element of L, so that $L = R(\vartheta)$. Order the conjugates of ϑ so that $\vartheta_1, \ldots, \vartheta_{\rho_1}$ are real and $\vartheta_{\rho_1+\nu}$ and $\vartheta_{\rho_1+\rho_2+\nu}$ are complex-conjugate for $\nu = 1, \ldots, \rho_2$, so that $\rho_1 + 2\rho_2 = nN$. Let ξ be a fixed one of the numbers $P_{j_1 \cdots j_m}(\alpha, \ldots, \alpha)$, where $j_1, \ldots, j_m$ satisfy the inequalities

$$0 \leq j_1 \leq r_1, \quad \cdots, \quad 0 \leq j_m \leq r_m, \quad \frac{j_1}{r_1} + \cdots + \frac{j_m}{r_m} \leq \mu. \tag{51}$$

Then ξ can be written as a polynomial in ϑ, with rational coefficients,

and as such has field conjugates $\xi^{(\nu)}$, $\nu = 1, \ldots, nN$. Hence we can define nN real numbers $\xi_1, \ldots, \xi_{nN}$ by the equations

$$\xi_\nu = \xi^{(\nu)}, \qquad \text{for } \nu = 1, \ldots, \rho_1,$$

$$\xi_\nu + i\xi_{\nu+\rho_2} = \xi^{(\nu)}, \qquad \text{for } \rho_1 + 1 \le \nu \le \rho_1 + \rho_2.$$

Collecting them in a fixed order for fixed coefficients $\gamma(s_1, \ldots, s_m)$ and for all $j_1, \ldots, j_m$ satisfying the inequalities (51), we have a set of numbers which can be considered as coordinates of a point; by Theorem 4-13 there are

$$M \le 2nNm^{\frac{1}{2}}\lambda^{-1}r$$

coordinates, and each is numerically smaller than $[b_2 N B_1^{1+3\delta}]+1=t$. Thus all the points, for the various sets of coefficients in C, lie in a cube of edge $2t$ in M-dimensional space. If each edge is divided into $3t$ equal parts, we get $(3t)^M$ subcubes of edge $\frac{2}{3}$. By (48), if

$$(1 + B_1)^{Nr} > (3t)^M, \tag{52}$$

there are more points than subcubes, and the points corresponding to two different polynomials $P^*(z_1, \ldots, z_m)$ and $P^{**}(z_1, \ldots, z_m)$ lie in the same subcube. If we put

$$\bar{P}(z_1, \ldots, z_m) = P^*(z_1, \ldots, z_m) - P^{**}(z_1, \ldots, z_m),$$

then

$$\overline{|\bar{P}_{j_1 \ldots j_m}(\alpha, \ldots, \alpha)|} \le \sqrt{2} \cdot \frac{2}{3} < 1$$

for $j_1, \ldots, j_m$ as in (51). Since $\bar{P}_{j_1 \ldots j_m}(\alpha, \ldots, \alpha)$ is an algebraic integer whose norm is numerically smaller than 1, it must be zero. Hence the index of $\bar{P}$ at the point $(\alpha, \ldots, \alpha)$ relative to $r_1, \ldots, r_m$ is at least μ. Also the coefficients $\bar{\gamma}(s_1, \ldots, s_m)$ in $\bar{P}$ are integers of K, not all zero, such that the relation (49) holds.

To verify (52), notice that by the inequality (40),

$$q_1^{\delta r_1} > 4b_2 N,$$

and hence

$$B_1 > 4b_2 N,$$

$$B_1^{Nr} > (4b_2 N B_1)^{\frac{1}{2}Nr},$$

$$B_1^{Nr} > (3b_2 N B_1^{1+3\delta} + 3)^{\frac{1}{2}Nr(1+3\delta)^{-1}},$$

$$(1 + B_1)^{Nr} > (3t)^M.$$

We now apply Theorem 4-12, the hypotheses of which are satisfied, as was noted earlier. Since $\bar{P}$ belongs to the class $\mathcal{R}_m(q_1^{\delta r_1}; r_1, \ldots, r_m)$, its index at $(\zeta_1, \ldots, \zeta_m)$ relative to $r_1, \ldots, r_m$ is less than η, defined in (45). Hence $\bar{P}$ possesses some derivative

$$Q(z_1, \ldots, z_m) = \frac{1}{k_1! \cdots k_m!} \left(\frac{\partial}{\partial z_1}\right)^{k_1} \cdots \left(\frac{\partial}{\partial z_m}\right)^{k_m} \bar{P},$$

with

$$\frac{k_1}{r_1} + \cdots + \frac{k_m}{r_m} < \eta,$$

such that

$$Q(\zeta_1, \ldots, \zeta_m) \neq 0.$$

The index of Q at the point $(\alpha, \ldots, \alpha)$ relative to $r_1, \ldots, r_m$ is at least $\mu - \eta$. Thus Q has the properties (a) and (b) of Theorem 4-14.

From the relations (49) and (50),

$$\lceil Q \rceil \leq 2^{r_1 + \cdots + r_m} b_2 N B_1 < 2^{m r_1} b_2 N B_1 < b_2 N B_1^{1+\delta}.$$

Hence for an arbitrary derivative,

$$\lceil Q_{i_1 \cdots i_m} \rceil \leq 2^{r_1 + \cdots + r_m} b_2 N B_1^{1+\delta} < b_2 N B_1^{1+2\delta}.$$

Finally,

$$\left|Q_{i_1 \cdots i_m}(z_1, \ldots, z_m)\right| < b_2 N B_1^{1+2\delta} \prod_{\nu=1}^{m} (1 + |z_\nu| + \cdots + |z_\nu|^{r_\nu})$$

$$< b_2 N B_1^{1+2\delta} \prod_{\nu=1}^{m} (1 + |z_\nu|)^{r_\nu}$$

$$< B_1^{1+3\delta} \prod_{\nu=1}^{m} (1 + |z_\nu|)^{r_\nu},$$

since $b_2 N < B_1^\delta$ by (40). The same inequality holds for the conjugate polynomials, and the proof is complete.

4-7 The Thue-Siegel-Roth theorem

THEOREM 4-15. *Let K be an algebraic number field of degree N, and let α be algebraic. Then for each $\varkappa > 2$, the inequality*

$$|\alpha - \zeta| < \frac{1}{(H(\zeta))^\varkappa} \tag{53}$$

has only finitely many solutions ζ in K.

Proof: We shall suppose that the theorem is false, so that (53) has infinitely many solutions, and produce a contradiction. We may suppose also that α is an integer. For if not there is a positive rational integer a such that $a\alpha$ is an algebraic integer, and for each solution ζ of (53) we have

$$|a\alpha - a\zeta| < \frac{a}{(H(\zeta))^{\varkappa}} \leq \frac{a^{\varkappa N+1}}{(H(a\zeta))^{\varkappa}}.$$

Hence for arbitrary $\epsilon > 0$, and for all solutions ζ with $H(\zeta)$ sufficiently large,

$$|a\alpha - a\zeta| < \frac{1}{(H(a\zeta))^{\varkappa - \epsilon}},$$

and ϵ can be chosen so small that $\varkappa - \epsilon > 2$.

Finally, it suffices to prove that (53) has only finitely many solutions in primitive elements ζ of K. For an algebraic number field has only finitely many subfields, and every element of K is a primitive element of some one of its subfields; moreover, the inequality in question does not depend on the degree of α over K.

We first choose m so large that $m > 4nm^{\frac{1}{2}}$ and

$$\frac{2m}{m - 4nm^{\frac{1}{2}}} < \varkappa, \tag{54}$$

which is possible since $\varkappa > 2$. For sufficiently small δ we have

$$m - 4(1 + 3\delta)nm^{\frac{1}{2}} - 2\eta > 0,$$

where η, given by (45), becomes arbitrarily small with δ. This condition is the same as that of (38). We choose δ to satisfy this and the inequality (37), and finally the inequality

$$\frac{2m(1 + \delta) + 2\delta N(2 + 5\delta)}{m - 4(1 + 3\delta)nm^{\frac{1}{2}} - 2\eta} < \varkappa, \tag{55}$$

which is possible in view of (54). The inequality (55) is equivalent to

$$\frac{m(1 + \delta) + \delta N(2 + 5\delta)}{\mu - \eta} < \varkappa, \tag{56}$$

by equations (43) and (44).

Having chosen m and δ, we now choose a solution ζ_1 of (53) (a primitive element of K) with $H(\zeta_1) = q_1$ and with q_1 so large as to

satisfy (40) and (42). We then choose further primitive solutions $\zeta_2, \ldots, \zeta_m$ of heights $q_2, \ldots, q_m$, such that for $j = 2, \ldots, m$,

$$\frac{\log q_j}{\log q_{j-1}} > \frac{2}{\delta}. \tag{57}$$

We now take r_1 to be any integer such that

$$r_1 > \frac{10 \log q_m}{\delta \log q_1}, \tag{58}$$

and define r_j, for $j = 2, \ldots, m$, by

$$\frac{r_1 \log q_1}{\log q_j} \leq r_j < \frac{r_1 \log q_1}{\log q_j} + 1. \tag{59}$$

Then the inequality (41) is satisfied. Also,

$$\frac{r_j \log q_j}{r_1 \log q_1} < 1 + \frac{\log q_j}{r_1 \log q_1} \leq 1 + \frac{\log q_m}{r_1 \log q_1} < 1 + \frac{\delta}{10}, \tag{60}$$

by (58). The conditions (39) are satisfied, since

$$r_m \geq \frac{r_1 \log q_1}{\log q_m} > 10\delta^{-1},$$

and

$$\frac{r_{j-1}}{r_j} > \frac{\log q_j}{\log q_{j-1}} \left(1 + \frac{\delta}{10}\right)^{-1} > \delta^{-1},$$

by (59), (60), and (57).

We know from Theorem 4–14 that there exists a polynomial $Q(z_1, \ldots, z_m)$, whose properties are listed in that theorem. Let $\zeta_1, \ldots, \zeta_m$ in K be zeros of irreducible polynomials of degree N with relatively prime coefficients in Z, the coefficients of z^N being $k_1, \ldots, k_m$, respectively. Then the number

$$\varphi = Q(\zeta_1, \ldots, \zeta_m)$$

is an element of K. If the field conjugates of ζ_i are $\zeta_i', \zeta_i'', \ldots$, for $i = 1, \ldots, m$, then $\mathbf{N}\varphi$ is a sum of products of powers of the $\zeta_i^{(j)}$ with integral coefficients from K, and in each such product a factor $\zeta_i^{(j)}$ occurs to the power r_i at most. In the proof of Theorem 2–21, it was shown that the product of k_i and any set of distinct conjugates of ζ_i is an algebraic integer. For each i, the field conjugates of ζ_i

are distinct, because ζ_i is a primitive element of K. It follows that $k_1{}^{r_1} \cdots k_m{}^{r_m} \mathbf{N}\varphi$ is an algebraic integer, and since it is also rational it is a rational integer, so that

$$|k_1{}^{r_1} \cdots k_m{}^{r_m} \mathbf{N}\varphi| \geq 1. \tag{61}$$

On the other hand, we have

$$Q(\zeta_1, \ldots, \zeta_m)$$
$$= \sum_{i_1=0}^{r_1} \cdots \sum_{i_m=0}^{r_m} Q_{i_1 \cdots i_m}(\alpha, \ldots, \alpha)(\zeta_1 - \alpha)^{i_1} \cdots (\zeta_m - \alpha)^{i_m},$$

and, by part (a) of Theorem 4-14, the terms with

$$\frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m} < \mu - \eta$$

all vanish. In all other terms we have

$$|(\zeta_1 - \alpha)^{i_1} \cdots (\zeta_m - \alpha)^{i_m}| < (q_1{}^{i_1} \cdots q_m{}^{i_m})^{-\varkappa}$$
$$= \{q_1{}^{i_1/r_1}(q_2{}^{r_2/r_1})^{i_2/r_2} \cdots (q_m{}^{r_m/r_1})^{i_m/r_m}\}^{-r_1\varkappa}$$
$$\leq (q_1{}^{i_1/r_1} \cdots q_1{}^{i_m/r_m})^{-r_1\varkappa}$$
$$< q_1{}^{-r_1(\mu-\eta)\varkappa},$$

since $q_j{}^{r_j/r_1} > q_1$ by (41). Hence, using part (c) of Theorem 4-14, we have

$$|\varphi| < (r_1 + 1) \cdots (r_m + 1)B_1{}^{1+3\delta}(1 + b_1)^{mr_1}q_1{}^{-r_1(\mu-\eta)\varkappa}$$
$$< B_1{}^{1+5\delta}q_1{}^{-r_1(\mu-\eta)\varkappa},$$

and by using part (c) again, together with Theorem 4-2, we obtain

$$|k_1{}^{r_1} \cdots k_m{}^{r_m} \mathbf{N}\varphi| < B_1{}^{1+5\delta}q_1{}^{-r_1(\mu-\eta)\varkappa}B_1{}^{(N-1)(1+3\delta)}$$
$$\times \prod_{i=1}^{m} \left\{ k_i \prod_{j=1}^{N} (1 + |\zeta_i{}^{(j)}|) \right\}^{r_i}$$
$$< B_1{}^{N(1+5\delta)}q_1{}^{-r_1(\mu-\eta)\varkappa} \prod_{i=1}^{m} (6^N q_i)^{r_i}.$$

Now, by (50),

$$6^{N(r_1+\cdots+r_m)} < 2^{3N(r_1+\cdots+r_m)} < B_1{}^{3\delta N} < q_1{}^{3\delta^2 N r_1} < q_1{}^{\delta N r_1},$$

so that

$$|k_1{}^{r_1} \cdots k_m{}^{r_m} \mathbf{N}\varphi| < q_1{}^{\delta N r_1(1+5\delta)+\delta N r_1+(r_1+\cdots+r_m)-r_1(\mu-\eta)\varkappa}$$
$$< q_1{}^{\delta N r_1(2+5\delta)+mr_1(1+\delta)-r_1(\mu-\eta)\varkappa}.$$

This, together with (61), implies that

$$\delta N(2 + 5\delta) + m(1 + \delta) > (\mu - \eta)\varkappa,$$

or

$$\varkappa < \frac{m(1 + \delta) + \delta N(2 + 5\delta)}{\mu - \eta},$$

which contradicts (56). This completes the proof.

4–8 Applications to Diophantine equations. The Thue-Siegel-Roth theorem will now be applied to show that a rather large variety of Diophantine equations have only finitely many solutions.

THEOREM 4–16. *Let $U(x, y)$ be a binary form of degree n, without multiple linear factors, whose coefficients belong to an algebraic number field K_0 of degree h. Let x and y be integral variables of K_0. Suppose that*

$$n > 2h.$$

Let $V(x, y)$ be any polynomial of total degree $\nu < n - 2h$ which has coefficients in K_0 and has no common factor with $U(x, y)$. Then the equation

$$U(x, y) = V(x, y) \tag{62}$$

has only finitely many solutions.

Proof: Just as in the representation theory for binary quadratic forms, it makes no difference whether we consider (62) or an equation obtained from it by a substitution $x = ax' + by'$, $y = cx' + dy'$, where a, b, c, d are in Z, and $|ad - bc| = 1$. If $U(x, y) = a_0x^n + \cdots + a_ny^n$, then

$$U(x, ax + y) = U(1, a)x^n + \cdots + a_ny^n,$$

$$U(x + by, y) = a_0x^n + \cdots + U(b, 1)y^n.$$

Choose a in K_0 so that $U(1, a) \neq 0$, and put $U(x, ax + y) = U_1(x, y)$. Then choose b in K_0 so that $U_1(b, 1) \neq 0$, and put $U_1(x + by, y) = U_2(x, y)$. Dropping the subscript, we see that there is no loss in generality in supposing that the coefficients of x^n and y^n in $U(x, y)$ are different from zero, and we can write

$$U(x, y) = \alpha y^n \prod_{k=1}^{n} \left(\frac{x}{y} - \xi_k\right), \tag{63}$$

where neither α nor any ξ_k is zero. By assumption, the numbers ξ_k are distinct, so if we put

$$c_1 = \min_{j \neq k} (|\xi_j - \xi_k|),$$

then $c_1 > 0$, and for every x and y, at least $n - 1$ of the factors in the product occurring in (63) have absolute values not less than $\frac{1}{2}c_1$.

Let $x = \eta$ and $y = \zeta \neq 0$ be integers of K_0, with field conjugates $\eta^{(1)}, \ldots, \eta^{(h)}, \zeta^{(1)}, \ldots, \zeta^{(h)}$. Then as we saw in Theorem 2–5,

$$\prod_{j=1}^{h} (\zeta^{(j)}t - \eta^{(j)}) = (Q(t))^f,$$

where $Q(t)$ is an irreducible polynomial with coefficients in Z, and $1 \leq f \leq h$. Let $M = \max (\lceil \zeta \rceil, \lceil \eta \rceil)$, and name the conjugates so that

$$Q(t) = \prod_{j=1}^{h/f} (\zeta^{(j)}t - \eta^{(j)}).$$

Then the coefficients of $Q(t)$ are numerically smaller then the corresponding coefficients of

$$\prod_{j=1}^{h/f} (Mt + M),$$

so that $\|Q\| \leq (2M)^{h/f}$. A fortiori, $H(\eta/\zeta) \leq (2M)^{h/f}$.

Now by Theorem 4–15, there are only finitely many solutions of the inequality

$$\left| \xi - \frac{\eta}{\zeta} \right| < \frac{1}{H(\eta/\zeta)^{2+\epsilon'}}$$

for fixed $\epsilon' > 0$. Hence for M sufficiently large, and $\epsilon = \epsilon'h$,

$$\left| \xi - \frac{\eta}{\zeta} \right| \geq \frac{1}{(2M)^{h(2+\epsilon')/f}} \geq \frac{1}{(2M)^{2h+\epsilon}},$$

at least if the left side is not zero. This is certainly true of the solutions of (62), since $U(x, y)$ and $V(x, y)$ have no common factor. The same argument applies to the numbers $\eta^{(j)}/\zeta^{(j)}$ and ξ_k, for $1 \leq j \leq h, 1 \leq k \leq n$; we see that for $\epsilon > 0$ and M sufficiently large, the inequality

$$\left| \xi_k - \frac{\eta^{(j)}}{\zeta^{(j)}} \right| > \frac{1}{(2M)^{2h+\epsilon}}, \qquad j = 1, \ldots, h; \quad k = 1, \ldots, n,$$

holds for every solution of (62). There is no loss in generality in sup-

posing that $M = |\zeta^{(1)}| = |\zeta|$, since (62) remains correct after replacing all quantities by their conjugates and if necessary interchanging x and y. Hence, for large M,

$$|U(\eta, \zeta)| > |\alpha| M^n \left(\frac{c_1}{2}\right)^{n-1} \frac{1}{(2M)^{2h+\epsilon}}.$$

On the other hand, there is a constant c_2, depending only on the coefficients of V, such that

$$|V(\eta, \zeta)| < c_2 M^\nu.$$

If we choose $\epsilon < n - 2h - \nu$, then for sufficiently large M,

$$|U(\eta, \zeta)| > |V(\eta, \zeta)|.$$

But a bound on M implies a bound on the integral coefficients of the polynomials defining η and ζ, so that there are only finitely many solutions of (62).

> COROLLARY. *If $U(x, y)$ is a binary form of degree $n > 2$, with coefficients in Z and without repeated linear factors, and if $a \neq 0$ is a rational integer, there are only finitely many rational integral solutions of the equation $U(x, y) = a$. In particular, the equation*
>
> $$ax^n + by^n = c$$
>
> *has only finitely many solutions in Z if a, b, and c are in Z, $abc \neq 0$, and $n \geq 3$.*

This follows immediately from the theorem, with $K_0 = R$, $h = 1$, and $n - 2h > 0$. The special case mentioned includes the higher-degree analog of Pell's equation, $x^n - dy^n = N$.

In the above considerations, strong use was made of the homogeneity of $U(x, y)$. If a Diophantine equation is not of the form specified in Theorem 4–16, it may still be possible to relate its solvability to that of one of this form. We now consider such a case.

4–9 A special equation. It was conjectured by E. Catalan in 1842 that 8 and 9 are the only two consecutive integers larger than 1 which are powers of other integers. This has never been proved; it has not even been shown that no three consecutive integers are powers, although it is trivial that no four can be, since one must be of the

form $4k + 2$. In slightly different terms, the problem is to show that the Diophantine equation

$$x^w - y^z = 1 \qquad (64)$$

has no solutions with w and z larger than 1, except for that mentioned. Various special cases arise by fixing, or specializing in some other way, one or more of the variables in (64). The case we are now going to examine is that in which the exponents are fixed, so that we consider the equation

$$x^m - y^n = 1. \qquad (65)$$

Catalan's conjecture would be proved if it could be shown that for each pair of integers m and n larger than 1, (65) has no positive solutions except that mentioned. Since this seems to be unfeasible, we consider the more modest question of whether (65) can have infinitely many solutions. This, at last, is a question that can be answered. It is a very weak consequence of the following theorem, due to Mahler, that (65) has only finitely many solutions if $m \geq 2$, $n \geq 3$.

THEOREM 4-17. *Suppose that $m \geq 2$, $n \geq 3$, $ab \neq 0$, $(x, y) = 1$. Then as* $\max (|x|, |y|) \to \infty$, *the greatest prime factor of*

$$ax^m + by^n$$

tends to infinity.

Since $x^2 - y^2 = 1$ has only the obvious solutions $x = \pm 1$, $y = 0$, the new problem is completely solved. Unfortunately Mahler's proof, which depends on a p-adic version of the Thue-Siegel theorem, cannot be included here. We can, however, obtain partial results of some interest.

If mn is even, the fact that (65) has only finitely many solutions is a consequence of the next theorem, which is a special case of a theorem proved anonymously and published by L. J. Mordell.

THEOREM 4-18. *Let $f(x)$ be a polynomial of degree $n \geq 3$, with coefficients in Z and with distinct zeros, and let a be any nonzero rational integer. Then the equation*

$$ay^2 = f(x) \qquad (66)$$

has only finitely many solutions x, y in Z.

Proof: Suppose that

$$f(x) = a_0(x - \xi_1) \cdots (x - \xi_n),$$

and that (66) has infinitely many solutions. The numbers $\alpha_j = a_0\xi_j$, for $j = 1, \ldots, n$, are algebraic integers, and if (66) holds, then

$$aa_0^{n-1}y^2 = (a_0x - \alpha_1) \cdots (a_0x - \alpha_n).$$

Let $K = R(\xi_1, \ldots, \xi_n)$ be the splitting field of f. Any ideal in K dividing $[a_0x - \alpha_i]$ and $[a_0x - \alpha_j]$ also divides $[\alpha_i - \alpha_j]$, so that the norm of such a common divisor is a divisor of the discriminant d of f. Hence, if P is a prime ideal divisor of y and $\mathbf{N}P > d$, then for some i, $P^2|[a_0x - \alpha_i]$. Since there are only finitely many ideals with norms smaller than d, and only finitely many divisors of $a_0^{n-1}a$, it follows that for each i,

$$[a_0x - \alpha_i] = B_iC_i^2, \tag{67}$$

where B_i and C_i are ideals, and B_i runs over a finite set of ideals.

Let D run over a fixed system of representatives of the various ideal classes in K; the number of D's is finite. Then for each i and some D, $C_i \sim D$, so that

$$[\beta]C_i = [\delta]D,$$

for some β and δ. We shall show that β can be chosen from a finite set of integers of K. Let $([\beta], [\delta]) = E$, and put $[\beta] = EF$, $[\delta] = EG$. Then $EFC_i = EDG$, whence $FC_i = DG$; thus $F|D$, and F is one of a finite set of ideals. By Theorem 3–2, there is an H with norm less than c (so that H is one of a finite set) such that $FH = [\gamma]$ is principal. Thus

$$[\gamma]C_i = (GH)D.$$

Since $C_i \sim D$, also $[\gamma] \sim GH$; hence $GH = [\zeta_i]$, and

$$[\gamma]C_i = [\zeta_i]D,$$

where γ is one of a finite set of integers.

By (67),

$$[\gamma^2][a_0x - \alpha_i] = B_i[\zeta_i^2]D^2,$$

from which it follows that B_iD^2 is principal, say $B_iD^2 = [\eta_i]$. Thus for some unit ϵ_i,

$$\gamma_i^2(a_0x - \alpha_i) = \epsilon_i\eta_i\zeta_i^2.$$

By Dirichlet's theorem on units, ϵ_i can be written as $\epsilon_i'\epsilon_i''^2$, where ϵ_i' is one of a finite number of units. Finally, for $i = 1, \ldots, n$,

$$a_0x - \alpha_i = \varkappa_i\lambda_i^2,$$

where $\lambda_1, \ldots, \lambda_n$ are integers of K, and $\varkappa_1, \ldots, \varkappa_n$ are certain ones of finitely many numbers of K. Hence

$$\varkappa_1 \lambda_1{}^2 - \varkappa_2 \lambda_2{}^2 = \alpha_2 - \alpha_1 \neq 0,$$

$$\varkappa_2 \lambda_2{}^2 - \varkappa_3 \lambda_3{}^2 = \alpha_3 - \alpha_2 \neq 0,$$

$$\varkappa_3 \lambda_3{}^2 - \varkappa_1 \lambda_1{}^2 = \alpha_1 - \alpha_3 \neq 0.$$

Now let $L = K(\sqrt{\varkappa_1}, \sqrt{\varkappa_2}, \sqrt{\varkappa_3})$. Then, in L,

$$(\lambda_1 \sqrt{\varkappa_1} - \lambda_2 \sqrt{\varkappa_2})(\lambda_1 \sqrt{\varkappa_1} + \lambda_2 \sqrt{\varkappa_2}) = \alpha_2 - \alpha_1,$$

and since the denominators of $\varkappa_1$ and $\varkappa_2$ can be taken to be bounded, it follows that

$$\lambda_1 \sqrt{\varkappa_1} - \lambda_2 \sqrt{\varkappa_2} = \beta_3 \epsilon_3{}^l,$$

where β_3 is one of finitely many elements of L, ϵ_3 is a unit of L, and $l > 1$ is an arbitrary positive integer. Similarly,

$$\lambda_2 \sqrt{\varkappa_2} - \lambda_3 \sqrt{\varkappa_3} = \beta_1 \epsilon_1{}^l,$$

$$\lambda_3 \sqrt{\varkappa_3} - \lambda_1 \sqrt{\varkappa_1} = \beta_2 \epsilon_2{}^l.$$

But then

$$\frac{\beta_1}{\beta_3}\left(\frac{\epsilon_1}{\epsilon_3}\right)^l + \frac{\beta_2}{\beta_3}\left(\frac{\epsilon_2}{\epsilon_3}\right)^l = -1. \tag{68}$$

If there were only finitely many distinct ratios ϵ_1/ϵ_3, there would be a finite set of coefficients φ such that

$$\sqrt{a_0 x - \alpha_2} - \sqrt{a_0 x - \alpha_3} = \varphi(\sqrt{a_0 x - \alpha_1} - \sqrt{a_0 x - \alpha_2})$$

for every solution x of (66) and for suitable determination of the radicals. This is clearly impossible, so (68) must have infinitely many solutions in integers ϵ_1/ϵ_3, ϵ_2/ϵ_3 of L. But for l sufficiently large, this is in contradiction with Theorem 4–16. Hence the supposition that (66) has infinitely many solutions is not tenable, and the proof is complete.

Returning to equation (65), we see that the only possible solutions have $x = 0$ or ± 1, if $(m, n) > 1$. For the problem that remains, it suffices to consider the case in which $m = p$ and $n = q$ are distinct odd primes. This was treated by M. Newman, whose work was not published. A slightly strengthened version of his result, obtained by

applying Theorem 4–16 rather than the analogous consequence of the Thue-Siegel theorem, follows.

THEOREM 4–19. *If p and q are distinct odd primes such that $q > 2(p - 1)$ and q does not divide the class number of the cyclotomic field $K_p = R(\zeta)$, where $\zeta = exp(2\pi i/p)$, then the equations*

$$x^p - y^q = \pm 1 \tag{69}$$

have only finitely many solutions x, y in Z.

Proof: We carry out the proof only for the equation $x^p - y^q = 1$; the alternate case requires only trivial modifications. Put $1 - \zeta = \pi$ and $[\pi] = P$, so that P is a prime ideal of K_p, by Theorem 3–6. Let h be the class number of K_p.

If x and y satisfy (69) with the plus sign, then

$$[x - 1][x - \zeta] \cdots [x - \zeta^{p-1}] = [y]^q. \tag{70}$$

Put

$$D_{rs} = [x - \zeta^r, x - \zeta^s] \qquad \text{for } 0 \le r \le p - 1,$$
$$0 \le s \le p - 1, \quad r \ne s.$$

Then

$$D_{rs} = [x - \zeta^r, \zeta^r - \zeta^s] = [x - 1 + 1 - \zeta^r, \zeta^r - \zeta^s]$$

$$= \left[x - 1 + \frac{1 - \zeta^r}{1 - \zeta}\pi, \frac{\zeta^r - \zeta^s}{1 - \zeta}\pi \right] = \left[x - 1 + \frac{1 - \zeta^r}{1 - \zeta}\pi, \pi \right]$$

$$= [x - 1, \pi],$$

since $(\zeta^k - \zeta^l)/(1 - \zeta)$ is a unit if $p \nmid (k - l)$. Thus D_{rs} is the same for all r and s, and, since $D_{rs} | P$ and P is prime, either $D_{rs} = [1]$ or $D_{rs} = P$. We consider the two cases separately.

If $D_{rs} = [1]$, then the ideals $[x - \zeta^r]$ are pairwise relatively prime; since their product is a qth power, there are ideals $A_0, \ldots, A_{p-1}$ such that

$$[x - \zeta^r] = A_r{}^q, \qquad r = 0, \ldots, p - 1. \tag{71}$$

Suppose that e_r is the smallest positive integer such that $A_r{}^{e_r}$ is principal; by Theorem 3–4, $e_r | h$, and by (71), $e_r | q$. But q is prime and $q \nmid h$, so $e_r = 1$ and A_r is principal. Hence there are integers α and β and

units ϵ and ϵ' of K_p such that $x - 1 = \epsilon\alpha^q$ and $x - \zeta = \epsilon'\beta^q$, whence

$$\epsilon'\beta^q - \epsilon\alpha^q = \pi. \tag{72}$$

By Theorem 2–45, the units of K_p have a finite basis, so that each unit has a representation $\epsilon_1 \cdot \epsilon_2^q$, where ϵ_1 is one of the finite number of units obtained by taking products of powers of the basis elements, with exponents non-negative and smaller than q. Thus (72) implies that one of the finitely many equations

$$\epsilon_1'(\epsilon_2'\beta)^q - \epsilon_1(\epsilon_2\alpha)^q = \pi \tag{73}$$

must hold. But for each choice of ϵ_1 and ϵ_1', (73) has only finitely many integral solutions $\epsilon_2\alpha$, $\epsilon_2'\beta$ in K_p; this is evident from Theorem 4–16 with $K_0 = K_p$, $h = p - 1$, $n = q > 2(p - 1)$, $\nu = 0$. Hence x, and therefore also y, has only finitely many possible values.

The proof for the case $D_{rs} = P$ proceeds similarly. We put $x - 1 = \pi w$ and $y = \pi^m z$, where w and z are integers of K_p with $[\pi, z] = [1]$. Then (70) becomes

$$[w]\left[w + \frac{1 - \zeta}{1 - \zeta}\right] \cdots \left[w + \frac{1 - \zeta^{p-1}}{1 - \zeta}\right] = P^{mq-p}[z]^q,$$

and since the ideals on the left are pairwise relatively prime, there is a t with $0 \le t \le p - 1$ such that

$$P^{mq-p}\Big|\left[w + \frac{1 - \zeta^t}{1 - \zeta}\right].$$

Thus there are ideals $A_0, \ldots, A_{p-1}$ such that

$$\left[w + \frac{1 - \zeta^t}{1 - \zeta}\right] = P^{mq-p}A_t^q,$$

$$\left[w + \frac{1 - \zeta^r}{1 - \zeta}\right] = A_r^q, \qquad \text{for } 0 \le r \le p - 1, \quad r \ne t.$$

As before, it follows that all the ideals A_r are principal (for $r = t$, use the fact that an ideal equivalent to a principal ideal is principal). Since $p > 2$, there are distinct rational integers r and s different from t such that $0 \le r \le p - 1$, $0 \le s \le p - 1$. Then for integers α and β and units ϵ and ϵ' of K_p,

$$w + \frac{1 - \zeta^r}{1 - \zeta} = \epsilon\alpha^q, \qquad w + \frac{1 - \zeta^s}{1 - \zeta} = \epsilon'\beta^q,$$

so that

$$\epsilon'\beta^q - \epsilon\alpha^q = \frac{\zeta^r - \zeta^s}{1 - \zeta},$$

and the expression on the right is not zero. The earlier reasoning shows that the theorem is also true in this case.

PROBLEMS

1. Extend Theorem 4–18 to the case that f may have multiple zeros, but has at least three distinct zeros of odd orders.

2. Deduce from the finiteness of the number of solutions of (66) that as the integral variable x tends to infinity, the greatest prime divisor of $f(x)$ does also. [*Hint:* Assume that for infinitely many x, $f(x)$ is a product of powers of a fixed finite set of primes, and obtain a contradiction.]

REFERENCES

Section 4–1

See the following papers: J. Liouville, *Journal des Mathématiques Pures et Appliquées* (Paris) **16**, 133–142 (1851); A. Thue, *Journal für die Reine und Angewandte Mathematik* (Berlin) **135**, 284–305 (1909); C. L. Siegel, *Mathematische Zeitschrift* (Berlin) **10**, 173–213 (1921); F. J. Dyson, *Acta Mathematica* (Stockholm) **79**, 225–240 (1947); T. Schneider, *Archiv der Mathematik* (Karlsruhe) **1**, 288–295 (1948–1949); K. F. Roth, *Mathematika* (London) **2**, 1–20 (1955); Corrigendum, *Mathematika* **2**, 168 (1955).

The paper by Siegel contains many variants and applications of the Thue-Siegel theorem.

Section 4–7

The literature concerning Catalan's conjecture is reviewed by R. Obláth, *Revista Matemática Hispano-Americana* (Madrid) **1**, 122–140 (1941). Mahler's theorem appeared in *Nieuw Archief voor Wiskunde* (Amsterdam) **1**, 113–122 (1953). Theorem 4–17 appeared in *Journal of the London Mathematical Society* **2**, 66–68 (1926).

CHAPTER 5

IRRATIONALITY AND TRANSCENDENCE

5–1 Irrational numbers. One of the oldest results in the theory of numbers is that $\sqrt{2}$ is irrational; this was known to the Pythagoreans in the fifth century B.C. The proof, when suitably generalized with the help of the Unique Factorization Theorem, leads to the well-known rule for determining the possible rational zeros of a polynomial with rational integral coefficients; this in turn makes it possible to show, if such is the case, that a given polynomial has only irrational zeros. Thus the numbers given implicitly as zeros of polynomials can be trivially classified as rational or irrational.

If a number is given by its decimal expansion, one has only to determine whether its digits eventually recur periodically to know whether or not it is irrational. For example, the number

$$0.1234567891011\ldots,$$

whose successive digits are formed in an obvious fashion, is clearly irrational, since arbitrarily long blocks of a single digit occur, precluding periodicity. Similarly, using the regular continued fraction expansion of a real number, one can identify not only the rational numbers but also the quadratic irrationalities. (Unfortunately, there is no simple algorithm known which singles out the algebraic numbers of fixed degree $n \geq 3$ in a distinctive way.)

If a real number x is not given in one of these convenient forms, the problem of deciding whether or not it is rational may be decidedly nontrivial. It is, for example, not known whether Euler's constant, defined as

$$\lim_{n \to \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \log n\right),$$

is rational. Aside from properties of special algorithms, the only method available for investigating such questions depends on the following observation. If $x = a/b$ is rational, then for every pair of integers p and q, the number $qx - p$ is some integral multiple of $1/b$,

so that it is impossible to find an infinite sequence of pairs p_n and q_n such that

$$|q_1x - p_1| > |q_2x - p_2| > |q_3x - p_3| > \cdots. \qquad (1)$$

More generally, no such sequence can be found for which

$$|q_nx - p_n| \neq 0 \text{ for every } n, \quad \text{and} \quad \lim_{n \to \infty} |q_nx - p_n| = 0. \qquad (2)$$

On the other hand, when x is irrational there are infinitely many solutions of the inequality

$$0 < |qx - p| < \frac{1}{q}.$$

We therefore have

THEOREM 5–1. *Each of the following is a necessary and sufficient condition for the irrationality of a real number x:*

(a) *there are integers $p_1, q_1, p_2, q_2, \ldots,$ such that the inequalities* (1) *hold;*

(b) *there are integers $p_1, q_1, p_2, q_2, \ldots,$ such that the conditions* (2) *hold.*

As a simple application of this principle, we prove

THEOREM 5–2. *The number e is irrational.*

Proof: We recall the expansion

$$\frac{1}{e} = 1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + \frac{(-1)^n}{n!} + \cdots.$$

It is well known that if $a_0, a_1, \ldots$ is an unbounded increasing sequence of positive numbers, then the series

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{a_k} \qquad (3)$$

converges to its sum S in such a way that

$$0 < \left| S - \sum_{k=0}^{n} \frac{(-1)^k}{a_k} \right| < \frac{1}{a_{n+1}}$$

for $n \geq 0$. Hence if we put $q_n = n!$ and

$$p_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!},$$

then p_n and q_n are integers and

$$0 < \left| q_n \cdot \frac{1}{e} - p_n \right| = n! \left| \frac{1}{e} - \sum_{k=0}^{n} \frac{(-1)^k}{k!} \right| < \frac{n!}{(n+1)!} = \frac{1}{n+1} .$$

It follows that $1/e$, and hence e itself, is irrational. (This is a variant of the original proof due to Fourier.) More generally, the same argument shows that if the LCM of the integers $a_1, \ldots, a_n$ is $o(a_{n+1})$ as $n \to \infty$, then the series (3) converges to an irrational number.

For completeness, we give a proof due to I. Niven that π is irrational. It is short and simple to follow, but to one unfamiliar with older work it must appear completely unmotivated.

THEOREM 5-3. *The number π is irrational.*

Proof: Suppose on the contrary that $\pi = a/b$, where a and b are integers. Put

$$f(x) = \frac{x^n (a - bx)^n}{n!} ,$$

and

$$F(x) = f(x) - f''(x) + f^{(iv)}(x) - \cdots + (-1)^n f^{(2n)}(x),$$

where the positive integer n will be specified later. Now $f(0) = f'(0) = \cdots = f^{(n-1)}(0) = 0$, and if we write

$$f(x) = \frac{a_0 x^n + a_1 x^{n+1} + \cdots + a_n x^{2n}}{n!} ,$$

we see that for $n \le k \le 2n$,

$$f^{(k)}(x) = \frac{1}{n!} \sum_{l=0}^{n} (n+l)(n+l-1) \cdots (n+l-k+1) a_l x^{n+l-k}$$

$$= \frac{1}{n!} \sum_{l=0}^{n} \frac{(n+l)!}{(n+l-k)!} a_l x^{n+l-k},$$

so that

$$f^{(k)}(0) = \frac{1}{n!} \cdot k! a_{k-n}.$$

Hence $f^{(j)}(0) \in Z$, and since $f(x) = f(\pi - x)$, also $f^{(j)}(\pi) \in Z$, for $0 \le j \le 2n$. Finally, $F(0)$ and $F(\pi)$ must be integers.

On the other hand,

$$\frac{d}{dx}\left(F'(x)\sin x - F(x)\cos x\right) = F''(x)\sin x + F(x)\sin x$$

$$= f(x)\sin x,$$

so that

$$\int_0^\pi f(x)\sin x\,dx = [F'(x)\sin x - F(x)\cos x]_0^\pi = F(\pi) + F(0).$$

But for $0 < x < \pi$,

$$0 < f(x)\sin x < \frac{\pi^n a^n}{n!},$$

so that the above integral is positive but arbitrarily small for n sufficiently large. But this is impossible, since $F(0) + F(\pi)$ is an integer. The contradiction establishes the theorem.

PROBLEM

Given a real number x, define the sequence $\{x_k\}$ of real numbers and the sequence $\{a_k\}$ of integers by the conditions

$$[x] = a_0, \qquad x_1 = x - [x],$$

$$x_1 = \frac{1}{a_1} + x_2, \qquad \text{where } \frac{1}{a_1} \le x_1 < \frac{1}{a_1 - 1},$$

$$x_2 = \frac{1}{a_2} + x_3, \qquad \text{where } \frac{1}{a_2} \le x_2 < \frac{1}{a_2 - 1},$$

$$\vdots$$

$$x_k = \frac{1}{a_k} + x_{k+1}, \qquad \text{where } \frac{1}{a_k} \le x_k < \frac{1}{a_k - 1},$$

$$\vdots$$

Thus

$$x = a_0 + \frac{1}{a_1} + \frac{1}{a_2} + \cdots.$$

Show that this expansion terminates if and only if x is rational. Show also that if x has an infinite series expansion

$$x = b_0 + \frac{1}{b_1} + \frac{1}{b_2} + \cdots.$$

where the numbers b_k are integers with $b_{k+1} > b_k{}^2$, then $b_k = a_k$ for all k, and x is irrational.

5-2 The existence of transcendental numbers. One class of irrationals, the algebraic numbers, has been treated in some detail in the preceding chapters. We now consider the complementary set of *transcendental* numbers: those complex numbers which do not satisfy any rational algebraic equation with coefficients in Z. It is by no means obvious that this set is nonvacuous; the first proof, given by Liouville in 1844, depends on the fact (see Theorem 4-1) that if α is algebraic of degree $n \geq 2$, then there is a constant C such that the inequality

$$\left| \alpha - \frac{p}{q} \right| < \frac{C}{q^n}$$

has no solution p, q in Z. If a number ξ can be found such that for every $\omega > 0$ the inequality

$$0 < |q\xi - p| < \frac{1}{q^\omega}, \qquad q > 1, \tag{4}$$

has a solution, then ξ cannot be algebraic of any degree, and must therefore be transcendental.

An example of a *Liouville number*, for which (4) always has a solution, is given by

$$\xi = \sum_{k=1}^{\infty} (-1)^k a^{-b_k},$$

where $a > 1$ is a fixed integer and $b_1, b_2, \ldots$ is an increasing sequence of positive integers such that

$$\limsup_{k \to \infty} \frac{b_{k+1}}{b_k} = \infty.$$

For, given ω, there is an $n = n(\omega)$ for which $b_{n+1}/b_n > \omega + 1$, and if we put

$$q = a^{b_n}, \qquad p = q \sum_{k=1}^{n} (-1)^k a^{-b_k},$$

then p and q are integers, and

$$0 < |q\xi - p| < q \cdot \frac{1}{a^{b_{n+1}}} = \frac{1}{q^{-1+b_{n+1}/b_n}} < \frac{1}{q^\omega}.$$

It should be emphasized that the condition (4), while sufficient for transcendence, is by no means necessary, even for real numbers.

For, using a modification of an argument due to Cantor, we can give a second proof of the existence of transcendental numbers, and in particular of numbers of this kind for which the inequality

$$\left| \xi - \frac{p}{q} \right| < \frac{1 - \epsilon}{3q^2}$$

has only finitely many solutions for fixed $\epsilon > 0$. It is known* that there are uncountably many irrational numbers ξ for which $M(\xi) = 3$, where $M(\xi)$ is the supremum of the numbers λ for which the inequality

$$\left| \xi - \frac{p}{q} \right| < \frac{1}{\lambda q^2}$$

has infinitely many solutions. Hence if the algebraic numbers are countable, it follows that there are nonalgebraic numbers for which $M(\xi) = 3$.

To order the algebraic numbers, we associate with each non-constant polynomial $P(x) = a_0 x^n + \cdots + a_n$ with integral coefficients the number $h(P) = n + |a_0| + \cdots + |a_n|$. There are no polynomials with $h(P) = 1$. If $h(P) = 2$, then $P(x) = x$ or $-x$. If $h(P) = 3$, then $P(x)$ is one of $\pm x \pm 1$, $\pm 2x$, $\pm x^2$, all combinations of signs being allowed. In general, it is clear that if $k \geq 2$, there are only finitely many polynomials such that $h(P) = k$. Hence all polynomials with integral coefficients can be arranged in a sequence: first those with $h(P) = 2$, in some order, then those with $h(P) = 3$, in some order, etc. Suppose that $P_1(x), P_2(x), \ldots$ is such a sequence. Each $P_k(x)$ has finitely many zeros; write down all the zeros of $P_1(x)$ in some order, then all those of $P_2(x)$ in some order, etc. Let this sequence be $\beta_1, \beta_2, \ldots$. Now if $\beta_2 = \beta_1$, delete β_2; if $\beta_3 = \beta_2$ or β_1, delete β_3; and in general, if β_k is equal to some β with smaller subscript, delete β_k. Then the resulting sequence $\alpha_1, \alpha_2, \ldots$ contains all algebraic numbers, each just once.

To summarize, if a number can be approximated sufficiently well by rational numbers, it is transcendental, but there are transcendental numbers which cannot be approximated even as well as some quadratic irrationalities.

---

* See, for example, Volume I, Theorem 9–12.

1. Show that ξ is a Liouville number if the partial quotients in its continued fraction expansion,

$$\xi = a_0 + \cfrac{1}{a_1 +} \cdots,$$

have the property that

$$\limsup_{k \to \infty} \frac{\log a_{k+1}}{\log ((a_1 + 1) \cdots (a_k + 1))} = \infty.$$

[*Hint:* Show from the recursion relation for the successive convergents that $q_k < (a_1 + 1) \cdots (a_k + 1)$, and then use Theorem 2–6.]

2. Investigate the implications of Theorem 4–15 as regards transcendental numbers.

5–3 A criterion for transcendence. In order to obtain an approximability condition which is equivalent to transcendence, we must replace the linear expresion $q\xi - p$ occurring in the inequality (4) by a polynomial in ξ.

THEOREM 5–4. *A real or complex number ξ is transcendental if and only if there corresponds to each $\omega > 0$ a positive integer n, such that the inequality*

$$0 < |x_0 + x_1\xi + \cdots + x_n\xi^n| < X^{-\omega} \tag{5}$$

has infinitely many integral solutions $x_0, \ldots, x_n$, where

$$X = \max (|x_0|, \ldots, |x_n|).$$

It is to be noticed that the Liouville numbers (those for which (4) has a solution for each ω) are precisely the numbers for which we can take $n = 1$ for every ω. In general, however, n increases with ω.

Proof: We first prove that the condition is sufficient. Let $\alpha = \alpha_1$ be algebraic of degree g, let $f(x) = a_0 + a_1x + \cdots + a_gx^g$ be that multiple of its defining polynomial which has relatively prime coefficients in Z, with $a_g > 0$, and let $\alpha_1, \ldots, \alpha_g$ be its conjugates. Let $h(x) = x_0 + x_1x + \cdots + x_nx^n$ $(x_n > 0)$ be any polynomial with integral coefficients, and with zeros $\beta_1, \ldots, \beta_n$ distinct from $\alpha_1, \ldots, \alpha_g$. Then

$$0 < \left| \frac{1}{a_g{}^n} \prod_{i=1}^{n} f(\beta_i) \right| = \left| \prod_{i=1}^{n} \prod_{j=1}^{g} (\beta_i - \alpha_j) \right| = \left| \prod_{j=1}^{g} \prod_{i=1}^{n} (\beta_i - \alpha_j) \right|$$

$$= \left| \frac{1}{x_n{}^g} \prod_{j=1}^{g} h(\alpha_j) \right|,$$

so that if $X = \max(|x_0|, \ldots, |x_n|)$, then

$$0 < |h(\alpha)| = \frac{\left| x_n{}^g \prod_{i=1}^{n} f(\beta_i) \right|}{a_g{}^n \prod_{j=2}^{g} |h(\alpha_j)|} = \frac{\left| x_n{}^g \prod_{i=1}^{n} f(\beta_i) \right|}{a_g{}^n \prod_{j=2}^{g} \left(\frac{|h(\alpha_j)|}{X} \right) \cdot X^{g-1}} \cdot \quad (6)$$

But

$$\prod_{i=1}^{n} f(\beta_i)$$

is a symmetric polynomial with integral coefficients in the β's, of degree g in each β, and is therefore, by the Symmetric Function Theorem, a polynomial of total degree g, with integral coefficients, in the elementary symmetric functions x_{n-1}/x_n, $-x_{n-2}/x_n, \ldots,$ $\pm x_0/x_n$. Hence the numerator in the expression (6) is a positive integer, and we have

$$|h(\alpha)| \geq \frac{1}{a_g{}^n \prod_{j=2}^{n} \left| \frac{h(\alpha_j)}{X} \right| \cdot X^{g-1}} \cdot$$

Now if $r = \lceil \alpha \rceil$, then

$$\left| \frac{h(\alpha_j)}{X} \right| \leq 1 + |\alpha_j| + |\alpha_j|^2 + \cdots + |\alpha_j|^n \leq 1 + r + r^2 + \cdots + r^n,$$

so that the quantity

$$\frac{1}{a_g{}^n \prod_{j=2}^{n} \left| \frac{h(\alpha_j)}{X} \right|}$$

has a positive lower bound $A(n, \alpha)$ depending only on α and n. Thus

$$|h(\alpha)| \geq \frac{A(n, \alpha)}{X^{g-1}} \cdot \quad (7)$$

It follows that if (5) has infinitely many solutions with ~~...~~ ξ cannot be algebraic of degree less than $\omega + 1$. Since ω can be arbitrarily large, ξ cannot be algebraic.

The necessity of the condition of Theorem 5–4 is a consequence of the following more general theorem.

THEOREM 5-5. *If $\vartheta_1, \ldots, \vartheta_n$ are complex numbers, then for a suitable c which depends only on n and $\vartheta_1, \ldots, \vartheta_n$, the inequality*

$$|x_0 + x_1\vartheta_1 + \cdots + x_n\vartheta_n| < \frac{c}{X^{\frac{1}{2}(n-1)}} \qquad (8)$$

has infinitely many integral solutions $x_0, \ldots, x_n$.

If ξ is transcendental, we can take $\vartheta_k = \xi^k$; since no polynomial in ξ vanishes, it follows that (5) has infinitely many solutions if $n = [2\omega + 2]$.

Proof: The theorem is trivial if $n = 1$. For $n > 1$, put

$$c' = c'(\vartheta_1, \ldots, \vartheta_n) = 1 + |\vartheta_1| + \cdots + |\vartheta_n|,$$

let $h \geq 2$ be a positive integer, and let $x_0', x_1', \ldots, x_n'$ range independently over the integers from $-h$ to h inclusive. Since each of the $n + 1$ numbers x_k' can assume any of $2h + 1$ values, there are $(2h + 1)^{n+1} = t$ expressions

$$L(\vartheta_1, \ldots, \vartheta_n) = x_0' + x_1'\vartheta_1 + \cdots + x_n'\vartheta_n, \qquad |x_k'| \leq h.$$

Let these be, in some order, $L_1, \ldots, L_t$. Clearly

$$|L_i(\vartheta_1, \ldots, \vartheta_n)| \leq c'h,$$

so that all the points $L_i(\vartheta_1, \ldots, \vartheta_n)$ lie in the square of side $2c'h$ with its center at the origin of the complex plane. Subdivide this square into m^2 subsquares of side $2c'h/m$ each; then if $m^2 < t$, there must be at least one subsquare containing more than one point $L(\vartheta_1, \ldots, \vartheta_n)$. We can fulfill the condition $m^2 < t$ by taking

$$m = [(2h + 1)^{\frac{1}{2}(n+1)}] - 1.$$

For this m, suppose that the points

$$L_1(\vartheta_1, \ldots, \vartheta_n) = x_0' + x_1'\vartheta_1 + \cdots + x_n'\vartheta_n$$

$$_?(\vartheta_1, \ldots, \vartheta_n) = \bar{x}_0' + \bar{x}_1'\vartheta_1 + \cdots + \bar{x}_n'\vartheta_n$$

lie in a co.   ubsquare; the distance between them does not exceed the length of the diagonal of the subsquare, which is $2\sqrt{2}\,c'h/m$. So if we put $x_0 = x_0' - \bar{x}_0', \ldots, x_n = x_n' - \bar{x}_n'$ (so that $X \leq h - (-h) = 2h$), and

$$L(\quad ..,\vartheta_n) = L_1(\vartheta_1, \ldots, \vartheta_n) - L_2(\vartheta_1, \ldots, \vartheta_n)$$

$$= x_0 + x_1\vartheta_1 + \cdots + x_n\vartheta_n,$$

then

$$|L(\vartheta_1, \ldots, \vartheta_n)| \leq \frac{2\sqrt{2}\,c'h}{[(2h+1)^{\frac{1}{2}(n+1)}] - 1} \leqq \frac{4c'h}{(2h)^{\frac{1}{2}(n+1)}}$$

$$= \frac{2c'}{(2h)^{\frac{1}{2}(n-1)}} \leqq \frac{2c'}{X^{\frac{1}{2}(n-1)}} \cdot \quad (9)$$

Hence (8) has at least one solution, with $c = 2c'$.

If $L(\vartheta_1, \ldots, \vartheta_n) = 0$, then $xL(\vartheta_1, \ldots, \vartheta_n) = L(x\vartheta_1, \ldots, x\vartheta_n) = 0$ for every integer x, and (8) has infinitely many solutions. In the contrary case, choose h_1 so large that

$$|L(\vartheta_1, \ldots, \vartheta_n)| > \frac{2c'}{(2h_1)^{\frac{1}{2}(n-1)}},$$

and repeat the entire argument with h replaced by h_1. Calling the new form thus produced $L^{(1)}$, we have, by the analog of (9) and the definition of h_1, that

$$|L^{(1)}(\vartheta_1, \ldots, \vartheta_n)| < |L(\vartheta_1, \ldots, \vartheta_n)|,$$

so that we have a second solution of (7). Continuing the process, we can obtain arbitrarily many solutions.

<div align="center">PROBLEM</div>

Show that if the numbers $\vartheta_1, \cdots, \vartheta_n$ are real, then Theorem 4–5 remains correct if the inequality (8) is replaced by

$$|x_0 + x_1\vartheta_1 + \cdots + x_n\vartheta_n| < \frac{c}{X^{n-1}} \cdot$$

5–4 Measure of transcendence. Mahler's classification. In light of Theorem 5–4, we make the following definition: a function $\varphi(n, t)$ is called a *transcendence measure* for the transcendental number ξ if for each n there is a constant c_n such that for every $X \geq 1$,

$$|x_0 + x_1\xi + \cdots + x_n\xi^n| > c_n\varphi(n, X)$$

for each set of integers $x_0, \ldots, x_n$ of height $X = \max(|x_0|, \ldots, |x_n|)$. By Theorem 5–5, any such $\varphi(n, t)$ is no larger than $t^{-\frac{1}{2}(n-1)}$. A theorem giving a measure of transcendence of a number ξ represents a refinement of the assertion that ξ is transcendental; such measures

have been given for certain numbers. In Section 5–5 we shall determine a measure of transcendence for e.

Mahler has elaborated on the theory of transcendence measure in the following way. Let z be a complex number, and put

$$\omega_n(X, z) = \omega_n(X) = \min \left(\left| \sum_{k=0}^{n} x_k z^k \right| \right),$$ (10)

where the minimum is extended over all those sets of rational integral coefficients $x_0, \ldots, x_n$ of heights at most X for which

$$\sum_{k=0}^{n} x_k z^k \neq 0.$$

Then $\omega_n(X)$ is at most 1, and is a nonincreasing function of both X and n. Put

$$\omega_n(X) = X^{-\rho_n(X)},$$ (11)

so that

$$\rho_n(X) = \frac{\log\left(1/\omega_n(X)\right)}{\log X},$$

and let

$$\omega_n(z) = \omega_n = \limsup_{X \to \infty} \rho_n(X),$$

$$\omega(z) = \omega = \limsup_{n \to \infty} \frac{\omega_n}{n}.$$

Each of ω_n and ω is either $+\infty$ or a non-negative number. If ω_n is infinite and $n' > n$, then $\omega_{n'}$ is also infinite; hence there is an index $\mu(z) = \mu$, which may be finite or infinite, such that ω_n is finite for $n < \mu$ and infinite for $n \geq \mu$. The two quantities ω, μ are never finite simultaneously, for the finiteness of μ implies that there is an $n < \infty$ such that $\omega_n = \infty$, whence $\omega = \infty$. The number z is called

an A-number,	if $\omega = 0$,	$\mu = \infty$,	
an S-number,	if $0 < \omega < \infty$,	$\mu = \infty$,	
a T-number,	if $\omega = \infty$,	$\mu = \infty$,	
a U-number,	if $\omega = \infty$,	$\mu < \infty$.	

If μ is finite, then there is a fixed integer n such that for every $\sigma > 0$ there are integers $x_0, \ldots, x_n$ such that

$$|x_0 + x_1 z + \cdots + x_n z^n| < X^{-\sigma}.$$

For the case $n = 1$, this is exactly the definition of the Liouville numbers, so that the U-numbers may be regarded as higher degree analogs of Liouville numbers. The author has shown that there are U-numbers of every degree.

If z is algebraic, the inequality (7) shows that $\rho_n(X)$, and hence also ω_n, remains bounded as $n \to \infty$, so that $\omega = 0$ and z is an A-number. If, on the other hand, z is transcendental, it follows from Theorem 5–5 that $\rho_n \geq \frac{1}{2}(n - 1)$, whence $\omega \geq \frac{1}{2}$. Thus the A-numbers are precisely the algebraic numbers.

The existence of T-numbers has never been proved.

THEOREM 5–6. *If the complex numbers z and w are algebraically dependent, that is, if there is a polynomial $F(x, y)$ with coefficients in Z such that $F(z, w) = 0$, then they belong to the same class.*

Proof: If z is algebraic and w is algebraically dependent on z, then w is clearly also algebraic. We may therefore suppose that z and w are transcendental.

Let
$$F(x, y) = \sum_{h=0}^{M} \sum_{k=0}^{N} a_{hk} x^h y^k,$$

and suppose that F is irreducible. (One consequence of this assumption is that no polynomial in x alone is a factor of F.) Write

$$F(x, y) = \sum_{h=0}^{M} A_h(y) x^h,$$

where
$$A_h(y) = \sum_{k=0}^{N} a_{hk} y^k.$$

We may suppose that $A_M(y)$ is not identically zero.

Let $A(x) = a_0 + \cdots + a_n x^n$ be a polynomial for which the minimum is achieved in the definition (10) of $\omega_n(X, z)$, so that in particular $\max(|a_k|) \leq X$. We shall obtain inequalities relating $\omega(z)$ and $\omega(w)$; since in the definition of these quantities the first limit is taken on X, we temporarily regard n as fixed and X as a parameter.

Since it is not the case that for each fixed y the polynomials $F(x, y)$ and $A(x)$ have a common zero, we know by a standard theorem* that

* See, for example, B. L. van der Waerden, *Modern Algebra* (English edition, translated by Fred Blum from the second revised German edition), New York: Frederick Ungar Publishing Co., 1949, Vol. 1, pp. 83–85.

the resultant

$$
R(y) = \begin{vmatrix}
a_0 & \cdots & \cdots & \cdots & a_n & 0 & \cdots & 0 \\
0 & a_0 & \cdots & \cdots & \cdots & a_n & 0 & \cdots \\
 & & \cdot & & & & \cdot & \\
 & & & \cdot & & & & \cdot \\
0 & \cdots & 0 & a_0 & \cdots & \cdots & \cdots & a_n \\
A_0(y) & \cdots & \cdots & \cdots & \cdots & A_M(y) & 0 & \cdots \\
 & \cdot & & & & & \cdot & \\
 & & \cdot & & & & & \cdot \\
\cdots & 0 & A_0(y) & \cdots & \cdots & \cdots & \cdots & A_M(y)
\end{vmatrix}
\begin{array}{l} \Big\} M \text{ rows} \\ \\ \\ \Big\} n \text{ rows} \end{array}
$$

is not identically zero. $R(y)$ is a polynomial in y of degree nN at most, with coefficients in Z. Since F is a fixed polynomial throughout, the coefficients in $R(y)$ do not exceed $c_1 X^M$, where c_1 is a constant depending only on n and F.

If for each l with $2 \le l \le M + n$, the lth column in the determinant for $R(y)$ is multiplied by x^{l-1} and added to the first column, the new first column is

$$
A(x), \ xA(x), \ \ldots, \ x^{M-1}A(x), \ F(x, y), \ xF(x, y), \ \ldots, \ x^{n-1}F(x, y).
$$

Expanding by minors of the new first column, we obtain an identity

$$
R(y) = A(x)g(x, y) + F(x, y)h(x, y),
$$

from which

$$
R(w) = A(z)g(z, w).
$$

Regarding $g(x, y)$ as a sum of minors, we see that its coefficients are rational integers not exceeding $c_2 X^{M-1}$ in absolute value, so that

$$
|g(z, w)| < c_3 X^{M-1}.
$$

Hence

$$
|A(z)| > c_3^{-1} X^{-M+1} |R(w)|.
$$

But

$$
|R(w)| \ge \omega_{nN}(c_1 X^M, w),
$$

so

$$
|A(z)| > c_3^{-1} X^{-M+1} \omega_{nN}(c_1 X^M, w).
$$

It follows from the definition of $A(x)$ that

$$\omega_n(X, z) \geq c_3^{-1} X^{-M+1} \omega_{nN}(c_1 X^M, w),$$

and so we obtain

$$\omega_n(z) = \limsup_X \frac{\log(1/\omega_n(X, z))}{\log X} \leq M - 1 + M\omega_{nN}(w),$$

$$\omega(z) = \limsup_n \frac{\omega_n(z)}{n}$$

$$\leq \limsup_n \frac{(M-1)N + MN\omega_{nN}(w)}{nN} \leq MN\omega(w),$$

and

$$\mu(w) \leq N\mu(z).$$

By symmetry,

$$\omega(w) \leq MN\omega(z) \quad \text{and} \quad \mu(z) \leq M\mu(w).$$

Thus $\omega(z)$ and $\omega(w)$ are simultaneously finite or infinite, as are $\mu(z)$ and $\mu(w)$; hence z and w are in the same class.

5–5 Arithmetic properties of the exponential function. In this section we shall prove a theorem due to Mahler which simultaneously shows that e is an S-number (and therefore transcendental), gives a transcendence measure for e, and shows that π is transcendental. The transcendence measure is not the most precise one known, but more exact results are more difficult to prove.

We begin with an algebraic analog of Theorem 5–5. Let $\omega_1, \ldots, \omega_m$ be distinct complex numbers (having no connection with the function $\omega_n(z)$ of the preceding section), and let $r_1, \ldots, r_m$ be positive integers. Instead of asking for rational integers $x_0, \ldots, x_m$ for which the quantity $x_0 + x_1\omega_1 + \cdots + x_m\omega_m$ is numerically small, we shall investigate the polynomials

$$A_k(z) = A_k(z; r_1, \ldots, r_m; \omega_1, \ldots, \omega_m), \qquad k = 1, \ldots, m,$$

of respective degrees $r_1 - 1, \ldots, r_m - 1$ at most, for which the function

$$R(z) = R(z; r_1, \ldots, r_m; \omega_1, \ldots, \omega_m)$$
$$= A_1(z)e^{\omega_1 z} + \cdots + A_m(z)e^{\omega_m z} \qquad (12)$$

is *algebraically* small, i.e., has a Maclaurin expansion beginning with a large power of z. The total number of coefficients among the

polynomials $A_k(z)$ is $r = r_1 + \cdots + r_m$; if they are taken as un-determined constants, then the conditions

$$R(0) = 0, \qquad R'(0) = 0, \qquad \ldots, \qquad R^{(r-2)}(0) = 0$$

yield a system of $r - 1$ linear homogeneous equations in these r unknowns. Such a system always has solutions distinct from $(0, 0, \ldots, 0)$. Let $R(z)$ temporarily designate any of the functions obtained in this manner; thus $R(z)$, which is not identically zero, certainly has a zero of order $r - 1$ at $z = 0$, and could conceivably have one of higher order there. Suppose that the actual order is $r - 1 + E$, so that $R(z)$ has an expansion

$$R(z) = \sum_{h=r+E-1}^{\infty} a_h z^h, \qquad a_{r+E-1} \neq 0.$$

The non-negative integer E is called the *excess*, and m is called the *order*, of $R(z)$. We first show that the excess is always equal to zero.

At least one of the polynomials $A_k(z)$ does not vanish identically, and with no loss in generality we may suppose it to be $A_1(z)$. It is easily proved by induction that if $D = d/dz$,

$$D^\alpha e^{\omega z} A(z) = e^{\omega z} (D + \omega)^\alpha A(z) \tag{13}$$

for every positive integer α and every function $A(z)$ with sufficiently many derivatives. Moreover, if $A(z)$ is a polynomial which is not identically zero, and $\omega \neq 0$, then $(D + \omega)^\alpha A(z)$ is a polynomial of the same degree as $A(z)$. Hence

$$D^{r_m} e^{-\omega_m z} R_m(z)$$
$$= D^{r_m} (A_1(z) e^{(\omega_1 - \omega_m)z} + \cdots + A_{m-1}(z) e^{(\omega_{m-1} - \omega_m)z} + A_m(z))$$
$$= A_1^*(z) e^{(\omega_1 - \omega_m)z} + \cdots + A^*_{m-1}(z) e^{(\omega_{m-1} - \omega_m)z}$$
$$= R(z; r_1, \ldots, r_{m-1}; \omega_1 - \omega_m, \ldots, \omega_{m-1} - \omega_m),$$

where A_1^* is not identically zero and, as implied by the notation, $\deg A_k^* \leq r_k - 1$ for $k = 1, \ldots, m - 1$. Clearly

$$R^{(\rho)}(0; r_1, \ldots, r_{m-1}; \omega_1 - \omega_m, \ldots, \omega_{m-1} - \omega_m) = 0$$

for $\rho = 0, 1, \ldots, r + E - r_m - 1$, so that from an R-function of order m and excess E we have obtained another of order $m - 1$ and excess E. Repeating the process, we come finally to a function $R_1(z) = R(z; r_1; \omega) = \tilde{A}(z) e^{\omega z}$ of order 1 and excess E. But if $\tilde{A}(z) = \tilde{a}_0 + \tilde{a}_1 z + \cdots + \tilde{a}_{r_1-1} z^{r_1-1}$, the conditions $R_1(0) = \cdots = R_1^{(r_1-2)}(0) = 0$ give $\tilde{a}_0 = \cdots = \tilde{a}_{r_1-2} = 0$, so that there is certainly

no such function which does not vanish identically, if $E > 0$. Hence $R^{(r-1)}(0) \neq 0$, or equivalently, the coefficient of z^{r-1} in the Maclaurin expansion of $R(z)$ is not zero, while all preceding coefficients are zero. Introducing an appropriate numerical factor, we can put

$$R(z) = \frac{z^{r-1}}{(r-1)!} + b_r z^r + \cdots.$$

The function R and the coefficients b_r, b_{r+1}, ... are now uniquely determined, since if there were two such functions for given $\omega_1, \ldots, \omega_m, r_1, \ldots, r_m$, their difference would have positive excess. Moreover, while we have so far known only that not all of the polynomials $A_1(z), \ldots, A_m(z)$ are identically zero, we now see that in fact they are of exact degrees $r_1 - 1, \ldots, r_m - 1$, respectively, since otherwise we could have begun with lower degree polynomials and arrived at a function of positive excess. Finally, we see that $R(z)$ is symmetric in the pairs of arguments $r_1, \omega_1; \cdots; r_m, \omega_m$, since the pairs can be permuted while the solution (subject to all the imposed conditions) is unique. This can also be seen by noting that $R(z)$ is the unique solution of the homogeneous linear differential equation

$$(D - \omega_1)^{r_1} \cdots (D - \omega_m)^{r_m} y = 0$$

for which $R(0) = \cdots = R^{(r-2)}(0) = 0$ and $R^{(r-1)}(0) = 1$, and the factors in the differential operator may be permuted at will.

We now obtain explicit expressions for $R(z)$ and the $A_k(z)$. Clearly

$$R(z; r_1; \omega_1) = \frac{z^{r_1-1}}{(r_1 - 1)!} e^{\omega_1 z}, \tag{14}$$

since this function has all the requisite properties and there is only one such function. Suppose that $R(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1})$ has already been determined. Then if J is the operator

$$J = \int_0^z \ldots dz,$$

we have by (13) that

$$(D-\omega_1)^{r_1} \cdots (D-\omega_\mu)^{r_\mu} \{ e^{\omega_\mu z} J^{r_\mu} (e^{-\omega_\mu z} R(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1})) \}$$

$$= (D-\omega_1)^{r_1} \cdots (D-\omega_{\mu-1})^{r_{\mu-1}}$$

$$\times \{ e^{\omega_\mu z} D^{r_\mu} J^{r_\mu} e^{-\omega_\mu z} R(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1}) \}$$

$$= (D-\omega_1)^{r_1} \cdots (D-\omega_{\mu-1})^{r_{\mu-1}} R(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1}) = 0,$$

and since $R(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1})$ has a zero of order $r_1 + \cdots + r_{\mu-1} - 1$ at 0, the function

$$e^{\omega_\mu z} J^{r_\mu}\left(e^{-\omega_\mu z} R(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1})\right)$$

has a zero of order $r_1 + \cdots + r_\mu - 1$ at 0, and it clearly has leading coefficient $((r_1 + \cdots + r_\mu - 1)!)^{-1}$. Hence

$$R(z; r_1, \ldots, r_\mu; \omega_1, \ldots, \omega_\mu)$$
$$= e^{\omega_\mu z} J^{r_\mu}\left(e^{-\omega_\mu z} R(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1})\right), \quad (15)$$

and consequently

$$R(z) = (e^{\omega_m z} J^{r_m})(e^{(\omega_{m-1}-\omega_m)z} J^{r_{m-1}}) \cdots (e^{(\omega_2-\omega_3)z} J^{r_2}) e^{(\omega_1-\omega_2)z} \frac{z^{r_1-1}}{(r_1-1)!}.$$

We now use the standard formula

$$J^\alpha f(z) = \int_0^z \frac{(z-t)^{\alpha-1}}{(\alpha-1)!} f(t)\, dt,$$

which is easily verified by integration by parts. We have

$$e^{(\omega_2-\omega_3)z} J^{r_2}\left(e^{(\omega_1-\omega_2)z} \frac{z^{r_1-1}}{(r_1-1)!}\right)$$

$$= e^{(\omega_2-\omega_3)z} \int_0^z \frac{t_1^{r_1-1}}{(r_1-1)!} \frac{(z-t_1)^{r_2-1}}{(r_2-1)!} e^{(\omega_1-\omega_2)t_1}\, dt_1$$

$$= \int_0^z \frac{t_1^{r_1-1}}{(r_1-1)!} \frac{(z-t_1)^{r_2-1}}{(r_2-1)!} e^{\omega_1 t_1 + \omega_2(z-t_1) - \omega_3 z}\, dt_1,$$

and by induction we see that

$$R(z) = \int_0^z dt_{m-1} \int_0^{t_{m-1}} dt_{m-2} \cdots$$

$$\cdots \int_0^{t_2}\left\{\frac{t_1^{r_1-1}(t_2-t_1)^{r_2-1} \cdots (t_{m-1}-t_{m-2})^{r_{m-1}-1}(z-t_{m-1})^{r_m-1}}{(r_1-1)!(r_2-1)! \cdots (r_{m-1}-1)!(r_m-1)!}\right.$$

$$\left. \times e^{\omega_1 t_1 + \omega_2(t_2-t_1) + \cdots + \omega_m(z-t_{m-1})}\right\} dt_1. \quad (16)$$

Before deducing an explicit formula for $A_k(z)$, we recall certain properties of inverse operators. The operator D^{-1}, as applied to an integral combination $f(z)$ of polynomials and exponential functions, yields that antiderivative which contains no constant of integration.

Hence

$$D^{-\rho}f(z) = J^{\rho}f(z) + \varphi(z),$$

where φ is a function annihilated by D^{ρ}, that is, a polynomial of degree $\rho - 1$ at most.. More generally, if $\omega \neq 0$ we define, by analogy to (13),

$$(D - \omega)^{-\rho}f(z) = e^{\omega z}D^{-\rho}\left(e^{-\omega z}f(z)\right),$$

so that

$$(D - \omega)^{-\rho}f(z) = e^{\omega z}J^{\rho}\left(e^{-\omega z}f(z)\right) + \psi(z), \qquad (17)$$

where $\psi(z)$ is annihilated by $(D - \omega)^{\rho}$; that is, it is $e^{\omega z}$ times a polynomial of degree $\rho - 1$ at most. Since

$$(D - \omega)\left(z^n + \frac{nz^{n-1}}{\omega} + \frac{n(n-1)z^{n-2}}{\omega^2} + \cdots + \frac{n!}{\omega^n}\right) = -\omega z^n,$$

and since no term of the operand is annihilated by $D - \omega$, we can write

$$(D - \omega)^{-1}z^n = -\sum_{r=0}^{n} \frac{n!}{r!\omega^{n-r+1}}z^r = -\frac{1}{\omega}\left(1 + \frac{D}{\omega} + \frac{D^2}{\omega^2} + \cdots\right)z^n.$$

More generally, it can be shown that if F is any polynomial of degree n for which $F(0) \neq 0$, then $(F(D))^{-1}z^n$ can be written as

$$(a_0 + a_1 D + \cdots + a_n D^n)z^n, \qquad (18)$$

where $a_0 + \cdots + a_n u^n$ is the Maclaurin expansion of $(F(u))^{-1}$ to $n + 1$ terms.*

We can now prove that for $k = 1, \ldots, m$,

$$A_k(z) = \left\{\prod_{\substack{h=1 \\ h \neq k}}^{m} (D + \omega_k - \omega_h)^{-r_h}\right\}\frac{z^{r_h-1}}{(r_k - 1)!}. \qquad (19)$$

For $m = 1$, the empty product is interpreted as the identity operator, of course, and in this case the correctness of (19) follows from equation (14). Suppose that it is correct for all polynomials

$$A_k(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1})$$

---

*A more complete discussion of inverse operators is given in E. L. Ince, *Ordinary Differential Equations*, New York: Longmans, Green & Co., Inc., 1926; reprinted by Dover Publications, New York, 1944; pp. 138–140.

with $\mu - 1$ pairs r_k, ω_k. Then, by (15) and (17),

$$R(z; r_1, \ldots, r_\mu; \omega_1, \ldots, \omega_\mu)$$

$$= e^{\omega_\mu z} J^{r_\mu} \sum_{k=1}^{\mu-1} A_k(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1}) e^{(\omega_k - \omega_\mu) z}$$

$$= e^{\omega_\mu z} \sum_{k=1}^{\mu-1} \{ e^{-(\omega_\mu - \omega_k) z} (D - \omega_\mu + \omega_k)^{-r_\mu}$$

$$\times A_k(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1}) + p_k(z) \}$$

$$= \sum_{k=1}^{\mu-1} e^{\omega_k z} (D - \omega_\mu + \omega_k)^{-r_\mu} A_k(z; r_1, \ldots, r_{\mu-1}; \omega_1, \ldots, \omega_{\mu-1}) + P(z) e^{\omega_\mu z},$$

where $p_k(z)$ and $P(z)$ are polynomials of degree $r_\mu - 1$ at most. It follows that (19) is correct for $k = 1$, for arbitrary m, and its truth for $k = 2$ follows from the previously noted symmetry of $R(z)$ in the pairs ω_k, r_k.

For fixed complex numbers $\omega_1, \ldots, \omega_m$, our considerations up to this point are valid for all the functions $R(z; r_1, \ldots, r_m; \omega_1, \ldots, \omega_m)$ corresponding to arbitrary sets $r_1, \ldots, r_m$ of positive integers. We now specialize the parameters so as to obtain a collection of functions depending on a single parameter ρ.

For h and k in the sequence $1, 2, \ldots, m$, define

$$\delta_{hk} = \begin{cases} 1 & \text{if } h = k, \\ 0 & \text{if } h \neq k, \end{cases}$$

and put

$$R_h(z) = R_h(z; \rho; \omega_1, \ldots, \omega_m) = R(z; \rho + \delta_{1h}, \ldots, \rho + \delta_{mh}; \omega_1, \ldots, \omega_m),$$

$$A_{hk}(z) = A_{hk}(z; \rho; \omega_1, \ldots, \omega_m) = A_k(z; \rho + \delta_{1h}, \ldots, \rho + \delta_{mh}; \omega_1, \ldots, \omega_m).$$

Here ρ is a fixed but arbitrary positive integer. We form the square matrix

$$A(z) = (A_{hk}(z)), \qquad h, k = 1, \ldots, m,$$

having determinant $D(z)$. Let the minor determinant of $A_{hk}(z)$ in $D(z)$ be $D_{hk}(z)$.

Now $A_{hk}(z)$ is a polynomial in z of degree $\rho + \delta_{hk} - 1$, and the coefficient of the highest power of z is, by (19),

$$\frac{1}{(\rho + \delta_{hk} - 1)!} \prod_{\substack{l=1 \\ l \neq k}}^{m} (\omega_k - \omega_l)^{-\rho - \delta_{hl}}.$$

Hence in the expansion of $D(z)$, the term formed from the elements of the main diagonal will be of higher degree than any other term, and $D(z)$ is therefore a polynomial of degree $m\rho$ with the coefficient of the highest power of z equal to

$$\frac{1}{(\rho!)^m} \prod_{k=1}^{m} \prod_{\substack{l=1 \\ l \neq k}}^{m} (\omega_k - \omega_l)^{-\rho}.$$

If, on the other hand, we solve the system of equations

$$\sum_{k=1}^{m} A_{hk}(z)e^{\omega_k z} = R_h(z), \qquad h = 1, \ldots, m,$$

for $e^{\omega_k z}$, we obtain the identity

$$D(z)e^{\omega_k z} = \sum_{h=1}^{m} (-1)^{h+k} D_{hk}(z)R_h(z).$$

Since the expansion of $R_h(z)$ begins with the term $z^{m\rho}/(m\rho)!$, the polynomial $D(z)$ is divisible by $z^{m\rho}$. Hence

$$D(z) = \frac{z^{m\rho}}{(\rho!)^m} \prod_{k=1}^{m} \prod_{\substack{l=1 \\ l \neq k}}^{m} (\omega_k - \omega_l)^{-\rho}, \tag{20}$$

and $D(z)$ vanishes only at $z = 0$.

Let $c_1, c_2, \ldots$ be positive constants depending only on $m, \omega_1, \ldots, \omega_m$. (In particular, they must not depend on ρ, which will eventually be large.)

Examination of equation (16) shows that for $1 \leq h \leq m$,

$$R_h(1) = O\left(\frac{c_1^{\rho}}{(\rho!)^m}\right).$$

From (19), we obtain

$$A_{hk}(z) = \left\{ \prod_{\substack{l=1 \\ l \neq k}}^{m} \sum_{\lambda_l=0}^{\infty} \binom{-\rho-\delta_{hl}}{\lambda_l} (\omega_k-\omega_l)^{-\rho-\delta_{hl}-\lambda_l} D^{\lambda_l} \right\} \frac{z^{\rho+\delta_{hk}-1}}{(\rho+\delta_{hk}-1)!},$$

where the sums need not be extended past the index ρ. Let

$$\Omega = \prod_{\substack{h,k=1 \\ h<k}}^{m} (\omega_k - \omega_h),$$

so that Ω can be regarded as a polynomial of total degree $m(m-1)/2$

in $\omega_1, \ldots, \omega_m$, with coefficients in Z not exceeding 2^{m^2} in absolute value. Since no exponent $\rho + \delta_{hl} + \lambda_l$ in the above sums exceeds $2\rho + 1$, the expression

$$a_{hk} = \Omega^{2\rho+1}\rho!A_{hk}(1)$$

is a polynomial in $\omega_1, \ldots, \omega_m$, of total degree

$$\frac{m(m-1)}{2}(2\rho+1)$$

at most, whose coefficients are rational integers of the order of magnitude $O(c_2^\rho \rho!)$. Finally, we put

$$r_h = \sum_{k=1}^{m} a_{hk}e^{\omega k}, \qquad h = 1, \ldots, m,$$

so that

$$r_h = \Omega^{2\rho+1}\rho!R_h(1) = O\left(\frac{c_3^\rho}{\rho!^{m-1}}\right).$$

The quantities $r_1, \ldots, r_m$ are linear forms in the numbers $e^{\omega k}$, and they are linearly independent, in the sense that no linear combination of the vectors $(a_{h1}, \ldots, a_{hm})$, for $1 \le h \le m$, is the zero vector. This is equivalent to the assertion that $D(1) \ne 0$, which follows from (20).

THEOREM 5–7. *Suppose that* $\omega_1, \ldots, \omega_m$ *all lie in an algebraic number field* K *of degree* g, *and let*

$$L_h = \sum_{k=1}^{m} b_{hk}e^{\omega k}, \qquad h = 1, \ldots, \mu,$$

be μ *independent linear forms in* $e^{\omega_1}, \ldots, e^{\omega_m}$ *with coefficients* b_{hk} *in* Z. *Suppose that*

$$m\left(1 - \frac{1}{g}\right) < \mu < m, \tag{21}$$

and put

$$b = \max_{\substack{1 \le h \le \mu \\ 1 \le k \le m}} (|b_{hk}|), \qquad s = \max_{1 \le h \le \mu} (|L_h|).$$

Then to each $\epsilon > 0$ *there corresponds a* $b_0(\epsilon)$ *such that* $s \ge b^{-\tau-\epsilon}$ *if* $b > b_0(\epsilon)$, *where*

$$\tau = \frac{m\mu g}{\mu g - m(g-1)} - 1.$$

Proof: By a well-known theorem on independence,* the μ forms $L_1, \ldots, L_\mu$, together with $m - \mu$ of the forms r_h, which we may designate by $r_{h_1}, \ldots, r_{h_{m-\mu}}$, are independent. Hence the determinant

$$\Delta = \begin{vmatrix} a_{h_1 1} & \cdots & a_{h_1 m} \\ \vdots & & \vdots \\ a_{h_{m-\mu} 1} & \cdots & a_{h_{m-\mu} m} \\ b_{11} & \cdots & b_{1m} \\ \vdots & & \vdots \\ b_{\mu 1} & \cdots & b_{\mu m} \end{vmatrix}$$

is not zero; it is obviously a polynomial in $\omega_1, \ldots, \omega_m$ of degree at most

$$\sigma = \frac{m(m-1)(2\rho+1)(m-\mu)}{2},$$

with coefficients in Z of the order of magnitude $O(c_4{}^\rho \rho!^{m-\mu} b^\mu)$. It follows, first, that there is a rational integer c_5 such that $c_5{}^\rho \Delta$ is an integer of K, and, second, that

$$\lceil \Delta \rceil = O\big((\sigma+1)^m c_4{}^\rho \rho!^{m-\mu} b^\mu c_6{}^\sigma\big)$$
$$= O(c_7{}^\rho \rho!^{m-\mu} b^\mu),$$

where c_6 is an upper bound for the various numbers $\lceil \omega_k \rceil$. Hence if $\Delta, \Delta'', \ldots, \Delta^{(g)}$ are the field conjugates of Δ, we have

$$\left| \frac{1}{\Delta} \right| = \left| \frac{c_5{}^\rho (c_5{}^\rho \Delta'') \cdots (c_5{}^\rho \Delta^{(g)})}{\mathbf{N}(c_5{}^\rho \Delta)} \right|$$
$$= O(c_8{}^\rho \rho!^{(g-1)(m-\mu)} b^{(g-1)\mu}).$$

Moreover, using subscripts on Δ to indicate minors, we have

$$\Delta_{lk} = O(c_9{}^\rho \rho!^{m-\mu-1} b^\mu), \qquad \Delta_{lk} r_{h_l} = O(c_{10}{}^\rho \rho!^{-\mu} b^\mu),$$
$$\text{for } 1 \leq l \leq m - \mu,$$

$$\Delta_{lk} = O(c_{11}{}^\rho \rho!^{m-\mu} b^{\mu-1}), \qquad \Delta_{lk} L_{l-m+\mu} = O(c_{12}{}^\rho \rho!^{m-\mu} b^{\mu-1} s),$$
$$\text{for } m - \mu + 1 \leq l \leq m.$$

---

* Cf. B. L. van der Waerden, *Modern Algebra* (English edition, translated by Fred Blum from the second revised German edition), New York: Frederick Ungar Publishing Co., 1949, Vol. 1, p. 101.

Using the identity

$$\Delta e^{\omega k} = \sum_{l=1}^{m-\mu} (-1)^{l+k} \Delta_{lk} r_{h_l} + \sum_{l=m-\mu+1}^{m} (-1)^{l+k} \Delta_{lk} L_{l-m+\mu},$$

it follows that

$$1 = O(c_{13}{}^{\rho} \rho!^{m(g-1)-\mu g} b^{\mu g}) + O(c_{14}{}^{\rho} \rho!^{(m-\mu)g} b^{\mu g-1} s),$$

or

$$c_{14}{}^{\rho} \rho!^{(m-\mu)g} b^{\mu g-1} s \geq c_{15} - c_{16} \cdot c_{13}{}^{\rho} \rho!^{m(g-1)-\mu g} b^{\mu g}. \qquad (22)$$

From the inequality (21), the exponent $m(g-1) - \mu g$ is negative. Hence the quantity

$$\frac{c_{13}{}^{\rho}}{\rho!^{\mu g - m(g-1)}}$$

may increase for small values of ρ, but it tends to zero as ρ increases indefinitely. At any rate, we can say that for b larger than some c_{17}, the smallest value of ρ for which

$$\frac{c_{15}}{2} \geq c_{16} c_{13}{}^{\rho} \rho!^{m(g-1)-\mu g} b^{\mu g}$$

is so large that $c_{13}{}^{\rho}$ is negligible as compared to the factorial, and for such ρ we have the asymptotic relation

$$\log \rho! \backsim \frac{\mu g}{\mu g - m(g-1)} \log b.$$

By (22), for $b > b_0(\epsilon)$,

$$s > b^{-\tau-\epsilon},$$

where

$$\tau = \mu g - 1 + \frac{(m-\mu)\mu g^2}{\mu g - m(g-1)} = \frac{m\mu g}{\mu g - m(g-1)} - 1.$$

This proves the theorem.

THEOREM 5-8. *Suppose that $\vartheta_1, \ldots, \vartheta_N$ are elements of an algebraic number field K of degree g, and that they are linearly independent over the rationals, so that no relation of the form*

$$d_1 \vartheta_1 + \cdots + d_N \vartheta_N = 0$$

holds with $d_1, \ldots, d_N$ rational and not all zero. Then if the coeffi-

cients $b_{\lambda_1 \cdots \lambda_N}$ *in the linear form*

$$L = \sum_{\lambda_1 = 0}^{M_1} \cdots \sum_{\lambda_N = 0}^{M_N} b_{\lambda_1 \cdots \lambda_N} e^{\lambda_1 \vartheta_1 + \cdots + \lambda_N \vartheta_N} \qquad (23)$$

in the quantities $e^{\lambda_1 \vartheta_1 + \cdots + \lambda_N \vartheta_N}$ *are rational integers with*

$$b = \max \left(|b_{\lambda_1 \cdots \lambda_N}| \right),$$

there is a constant T, *depending only on* g *and* N, *such that for sufficiently large* b,

$$|L| \geq b^{-T M_1 \cdots M_N}.$$

Proof: Let $\mu_1, \ldots, \mu_N$ be positive integers, and consider the quantities

$$L_{l_1 \cdots l_N} = e^{l_1 \vartheta_1 + \cdots + l_N \vartheta_N} L,$$

$$l_1 = 0, 1, \ldots, \mu_1; \quad \ldots; \quad l_N = 0, 1, \ldots, \mu_N, \qquad (24)$$

their number being

$$\mu = (\mu_1 + 1) \cdots (\mu_N + 1).$$

If we introduce the exponential factor in (24) inside the summation in (23); we see that the various $L_{l_1 \cdots l_N}$ may be regarded as linear forms in the quantities

$$e^{\omega_{\lambda_1 \cdots \lambda_N}},$$

where

$$\omega_{\lambda_1 \cdots \lambda_N} = \lambda_1 \vartheta_1 + \cdots + \lambda_N \vartheta_N,$$

$$\lambda_1 = 0, 1, \ldots, M_1 + \mu_1; \quad \ldots; \quad \lambda_N = 0, 1, \ldots, M_N + \mu_N,$$

the number of ω's being

$$m = (M_1 + \mu_1 + 1) \cdots (M_N + \mu_N + 1).$$

The numbers $\omega_{\lambda_1 \cdots \lambda_N}$ are distinct on account of the independence of $\vartheta_1, \ldots, \vartheta_N$ over the rationals, so that we can speak of the independence of the forms $L_{l_1 \cdots l_N}$. To see that as a matter of fact they are independent, order the subscript sets $\lambda_1 \ldots \lambda_N$ and $l_1 \ldots l_N$ by interpreting the λ's and l's as digits in the base q, for some sufficiently large q. Then there cannot be a linear relation among the coefficient vectors of any set of forms, since the $\omega_{\lambda_1 \cdots \lambda_N}$ with largest subscript occurs only in the form $L_{l_1 \cdots l_N}$ with largest subscript. Finally, there are positive constants α and β which are independent of the coeffi-

cients $b_{\lambda_1 \cdots \lambda_N}$, for which

$$\alpha < \left| \frac{L}{L_{l_1 \cdots l_N}} \right| < \beta.$$

It follows from Theorem 5–7 that if

$$\frac{m}{\mu} = \prod_{i=1}^{N} \frac{M_i + \mu_i + 1}{\mu_i + 1} < \frac{g}{g-1} , \qquad (25)$$

then for $b > b(\epsilon)$,

$$|L| \geq b^{-\tau - \epsilon},$$

where

$$\tau = \frac{g(M_1 + \mu_1 + 1) \cdots (M_N + \mu_N + 1)(\mu_1 + 1) \cdots (\mu_N + 1)}{g(\mu_1 + 1) \cdots (\mu_N + 1) - (g-1)(M_1 + \mu_1 + 1) \cdots (M_N + \mu_N + 1)} - 1.$$

Condition (25) is satisfied if

$$\mu_i = \left[\frac{M_i}{\left(\dfrac{2g}{2g-1} \right)^{1/N} - 1} \right],$$

since then

$$1 + \frac{M_i}{\mu_i + 1} \leq \left(\frac{2g}{2g-1} \right)^{1/N} ,$$

$$\prod_{i=1}^{N} \left(1 + \frac{M_i}{\mu_i + 1} \right) \leq \frac{2g}{2g-1} < \frac{g}{g-1} .$$

With this choice of μ_i we have

$$\tau = \frac{\mu \prod_{i=1}^{N} \left(1 + \dfrac{M_i}{\mu_i + 1} \right)}{1 - \dfrac{g-1}{g} \prod_{i=1}^{N} \left(1 + \dfrac{M_i}{\mu_i + 1} \right)} - 1 \leq \mu \frac{\dfrac{2g}{2g-1}}{1 - \dfrac{g-1}{g} \dfrac{2g}{2g-1}} - 1$$

$$= 2g\mu - 1. \qquad (26)$$

Since $g \geq 1$ and $N \geq 1$, we have $\mu_i \geq M_i \geq 1$. Since $[x] + 1 \leq 2x$ for $x \geq 1$, we have

$$\mu \leq \prod_{i=1}^{N} \frac{2M_i}{\left(\dfrac{2g}{2g-1} \right)^{1/N} - 1} ,$$

and we have the theorem with

$$T = \frac{2^{N+1}g}{\prod\limits_{i=1}^{N} \left\{ \left(\dfrac{2g}{2g-1} \right)^{1/N} - 1 \right\}}.$$

Taking $N = 1$, we have

COROLLARY 1.   *If $\vartheta \neq 0$ is algebraic, e^ϑ is an S-number, and in particular is transcendental.*

For $\vartheta = \pi i$, $e^\vartheta = -1$ is not transcendental.   Hence

COROLLARY 2.   *π is transcendental.*

We also have the following result, first proved by F. Lindemann in 1882.

COROLLARY 3.   *If $\vartheta_1, \ldots, \vartheta_N$ are algebraic and are linearly independent over the rationals, then $e^{\vartheta_1}, \ldots, e^{\vartheta_N}$ are algebraically independent over the field of algebraic numbers, that is, there is no polynomial $P(z_1, \ldots, z_N)$ with algebraic coefficients not all zero for which*

$$P(e^{\vartheta_1}, \ldots, e^{\vartheta_N}) = 0.$$

Finally, for $N = 1$ the brackets can be omitted in the definition of μ_1; then $\mu_1 = (2g - 1)M_1$ and

$$\tau \leq 2g\mu - 1 \leq 2g\left((2g-1)M_1 + 1 \right) - 1 = 2g(2g-1)M_1 + 2g - 1.$$

COROLLARY 4.   *If $\vartheta \neq 0$ is algebraic of degree g, then the function*

$$\varphi(n, t) = t^{-2g(2g-1)n-2g+1}$$

is a transcendence measure for e^ϑ.

5–6 A theorem of Schneider.   In addition to the Liouville numbers and values of the exponential function, many other specific numbers are known to be transcendental.   To indicate the type of results known, we mention the following:

(a) The Bessel functions $J_0(x)$ and $J_0'(x)$ are transcendental for algebraic $x \neq 0$.

(b) If α and β are algebraic, $\alpha \neq 0$ or 1, and β is irrational, then α^β is transcendental.   (In particular, $e^\pi = (-1)^{-i}$ is included.)

(c) At least one of the numbers $g_2, g_3, \omega_1, \omega_2$ associated with a Weierstrass $\wp$-function is transcendental, and if g_2 and g_3 are algebraic, at least one of z and $\wp(z)$ is transcendental.

(d) If $f(x)$ is a polynomial whose value is in Z for argument in Z, and $f(x) > 0$ for $x > 0$, then the number

$$0.\, f(1)f(2)f(3) \ldots ,$$

formed by juxtaposing the decimal representations of the values $f(x)$, is transcendental. (An example is the number $0.1361015 \ldots$, generated by $f(x) = (x^2 + x)/2$.)

(e) If ω is a positive quadratic irrationality, then the number

$$\sum_{n=0}^{\infty} [n\omega]z^n$$

is transcendental for algebraic $z \neq 0$.

On the other hand, it is not known whether the following numbers are transcendental:

(a) $\displaystyle \gamma = \lim_{n \to \infty} \left(1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n \right),$

(b) $\displaystyle \zeta(2n+1) = \sum_{k=1}^{\infty} \frac{1}{k^{2n+1}},$

(c) $\Gamma(x)$ for algebraic x not in Z,

(d) $e + \pi$, $\quad e\pi$.

The methods used to prove what little is known about specific transcendental numbers show considerable variety, both in technique and conception. T. Schneider has recently shown, however, that several results which earlier required separate proofs can all be obtained from a single theorem. This theorem says nothing directly about transcendental numbers; rather, its sense is that if several transcendental functions assume algebraic values at a large number of points, then they must either have large rates of growth or be algebraically dependent (as functions). The prototype of Schneider's result, proved by G. Pólya in 1920, asserts that if f is an integral transcendental function which assumes values in Z for $z = 0, 1, 2, \ldots,$ then

$$\limsup_{r \to \infty} \frac{M(r)}{2^r} \geq 1,$$

where, as usual,

$$M(r) = \max_{|z|=r}(|f(z)|).$$

There have been many refinements and extensions of Pólya's work, of course; we mention only that by A. Gelfond in 1929, where this kind of theorem was first used for transcendence investigations. (His result was that α^β is transcendental for algebraic $\alpha \neq 0, 1$, if β is an imaginary quadratic irrationality.)

In this section we shall prove Schneider's theorem, and in the next we shall apply it to the numbers α^β. (The facts mentioned above concerning the $\wp$-function can also be deduced, but the requisite preliminaries preclude doing so here.) Since the statement of the theorem is complicated, we first introduce some notation.

By the *order* of an entire function $f(z)$ we mean, as usual, the quantity

$$\limsup_{R\to\infty} \frac{\log\log M(R)}{\log R};$$

if $f(z)$ is of order μ, then

$$f(z) = O(e^{R^{\mu+\epsilon}})$$

as $|z| = R \to \infty$, for every fixed $\epsilon > 0$. Let $\zeta_1, \zeta_2, \ldots$ be an infinite sequence of complex numbers. Designate by

$$z_0(m) = z_0, \quad \ldots, \quad z_k(m) = z_k$$

the distinct numbers among $\zeta_1, \ldots, \zeta_m$, and by $l_\varkappa(m) + 1 = l_\varkappa + 1$ the multiplicity of occurrence of $z_\varkappa$ among $\zeta_1, \ldots, \zeta_m$. Thus

$$\sum_{\varkappa=0}^{k} (l_\varkappa + 1) = m. \tag{27}$$

Let $r(m) = r$ be the radius of the smallest circle about the origin which contains $z_1, \ldots, z_k$, and put

$$\alpha = \liminf_{m\to\infty} \frac{\log m}{\log r}, \tag{28}$$

so that $\alpha \leq \infty$. Let

$$l = \max(l_0, \ldots, l_k).$$

Finally, let K be a fixed algebraic number field of degree g, and, as

always, let $\lceil a \rceil$ be the maximum of the absolute values of the conjugates of a, for a in K.

THEOREM 5–9.   *Let* $f_1(z), \ldots, f_n(z)$ *be meromorphic functions with the property that for each* m, *the numbers*

$$f_\nu^{(\lambda)}(z_\varkappa), \qquad \lambda = 0, \ldots, l_\varkappa, \quad \varkappa = 0, \ldots, k, \quad \nu = 1, \ldots, n,$$

are in K. *Let* $H_\nu(z_\varkappa)$ *be positive rational integers such that all the numbers*

$$H_\nu(z_\varkappa)f_\nu^{(\lambda)}(z_\varkappa), \qquad \lambda = 0, \ldots, l_\varkappa, \quad \varkappa = 0, \ldots, k, \quad \nu = 1, \ldots, n,$$

are integers in K. *Suppose that*

$$l \leq \frac{m}{\log m}. \tag{29}$$

For each ν, *if* $f_\nu(z)$ *is entire let it be of order* μ_ν, *and otherwise suppose that there is an entire function* $G_\nu(z)$ *of order* μ_ν *such that* $G_\nu(z)f_\nu(z)$ *is entire and also of order* μ_ν. *Suppose that*

$$\frac{\mu_1 + \cdots + \mu_n}{n - 1} < \alpha, \tag{30}$$

and put

$$\eta_\nu = \frac{\mu_\nu}{\alpha}, \qquad \nu = 1, \ldots, n.$$

Suppose finally that

$$\limsup_{m \to \infty} \frac{\log \log \max_{0 \leq \varkappa \leq k} \left(|G_\nu(z_\varkappa)|^{-1} \right)}{\log m} \leq \eta_\nu, \qquad \nu = 1, \ldots, n, \tag{31}$$

and

$$\limsup_{m \to \infty} \frac{\log \log \max_{\substack{0 \leq \varkappa \leq k \\ 0 \leq \lambda \leq l_\varkappa}} \left(|f_\nu^{(\lambda)}(z_\varkappa)|, \; H_\nu(z_\varkappa) \right)}{\log m} \leq \eta_\nu, \qquad \nu = 1, \ldots, n. \tag{32}$$

Then $f_1, \ldots, f_n$ *are algebraically dependent over* K.

Proof: We form a polynomial

$$\Phi(z) = \sum_{\tau_1 = 0}^{t_1} \cdots \sum_{\tau_n = 0}^{t_n} C_{\tau_1 \cdots \tau_n} f_1^{\tau_1}(z) \cdots f_n^{\tau_n}(z),$$

and seek to determine the coefficients $C_{\tau_1 \cdots \tau_n}$ so that Φ has a zero of order $l_x + 1$ at z_x, for $x = 1, \ldots, k$. Here the numbers k, z_x, and l_x are all defined in terms of the sequence $\zeta_1, \zeta_2, \ldots$ and an index m, as explained earlier; m is fixed, and will be specified more exactly later. The conditions imposed on Φ require that all the numbers

$$\Phi^{(\lambda)}(z_x), \qquad \lambda = 0, \ldots, l_x; \quad x = 0, \ldots, k,$$

shall vanish, and this in turn yields a set of m homogeneous linear equations of the form

$$\sum_{\tau} \omega_\mu C_{\tau_1 \cdots \tau_n} = 0, \qquad \mu = 1, \ldots, m, \tag{33}$$

in the $t = (t_1 + 1) \cdots (t_n + 1)$ unknowns $C_{\tau_1 \cdots \tau_n}$. (Of course, the numbers ω_μ also depend on $\tau_1, \ldots, \tau_n$.) We put

$$t_\nu = [(2m^{1+\eta_1+\cdots+\eta_n-n\eta_\nu})^{1/n}], \qquad \nu = 1, \ldots, n, \tag{34}$$

so that

$$t = (t_1 + 1) \cdots (t_n + 1) \geq \left\{ \prod_{\nu=1}^{n} 2m^{1+\eta_1+\cdots+\eta_n-n\eta_\nu} \right\}^{1/n} = 2m. \tag{35}$$

The coefficients ω_μ in equations (33) are by assumption numbers in K, and after multiplication by the rational integral factor

$$\prod_{\nu=1}^{n} (H_\nu(z_x))^{t_\nu}$$

they become integers, say Ω_μ, of K. The size of the coefficients Ω_μ is determined in part by this numerical factor, in part by the values of the f_ν and their derivatives at the various points z_x, and in part by the numerical coefficients introduced by differentiation. In the estimate (36) below, the second of these is accounted for by (32). The third depends only on the set of exponents $t_1, \ldots, t_n$ and the order of the derivative considered, and so can be computed from the fact that the sum of all the coefficients in the expansion of

$$\frac{d^\lambda}{dz^\lambda} (z^{\tau_1} \cdots z^{\tau_n})$$

by the product formula is

$$\sum_{\nu=1}^{n} \tau_\nu \left(\sum_{\nu=1}^{n} \tau_\nu - 1 \right) \cdots \left(\sum_{\nu=1}^{n} \tau_\nu - \lambda + 1 \right) \leq \left(\sum_{\nu=1}^{n} \tau_\nu \right)^\lambda.$$

We thus obtain the bound

$$\lceil \Omega_\mu \rceil \leq \prod_{\nu=1}^{n} (H_\nu(z_\kappa))^{t_\nu} \cdot \prod_{\nu=1}^{n} \exp\left(t_\nu m^{\eta_\nu + \epsilon_\nu}\right) \cdot \left(\sum_{\nu=1}^{n} t_\nu\right)^l, \qquad (36)$$

where $\epsilon_\nu > 0$ and $\epsilon_\nu \to 0$ as $m \to \infty$. (Hereafter we designate any quantity with the latter properties by ϵ, and any positive integer independent of m, l, and r by γ.)

It follows from the inequality (30) and the definition of η_ν that

$$\eta_1 + \cdots + \eta_n < n - 1, \qquad (37)$$

so that, by (34),

$$\left(\sum_{\nu=1}^{n} t_\nu\right)^l < (\gamma m)^l.$$

Using this, together with (32) and (36), we obtain

$$\lceil \Omega_\mu \rceil < (\gamma m)^l \exp\left\{2 \sum_{\nu=1}^{n} 2^{1/n} m^{(1 + \eta_1 + \cdots + \eta_n - n\eta_\nu)/n + \eta_\nu + \epsilon}\right\}$$

$$< \gamma^m m^l \gamma^{m(1 + \eta_1 + \cdots + \eta_n)/n + \epsilon}.$$

By (29), $m^l < \gamma^m$. By (37), $\eta_1 + \cdots + \eta_n = n - 1 - \delta$ with δ a positive constant; hence

$$\frac{1}{n}(1 + \eta_1 + \cdots + \eta_n) + \epsilon = \frac{n - \delta}{n} + \epsilon = 1 - \frac{\delta}{n} + \epsilon.$$

We henceforth require m to be so large that

$$\epsilon < \frac{\delta}{n}. \qquad (38)$$

Then $$\lceil \Omega_\mu \rceil < \gamma^m. \qquad (39)$$

Using this and (35), we shall now show that *there are coefficients* $C_{\tau_1 \cdots \tau_n}$ *satisfying* (33) *which are integers in* K, *are not all zero, and are such that*

$$\lceil C_{\tau_1 \cdots \tau_n} \rceil < \gamma^m. \qquad (40)$$

To simplify the notation, arrange the $C_{\tau_1 \cdots \tau_n}$ in some fixed linear order, and rewrite (33) in the form

$$\sum_{\tau=1}^{t} \Omega_{\mu\tau} C_\tau = 0, \qquad \mu = 1, \ldots, m.$$

Let $\rho_1, \ldots, \rho_g$ be an integral basis for K, let $h \in Z$ be positive, and

let B be the set of integers in K of the form

$$b_1\rho_1 + \cdots + b_g\rho_g,$$

where the b's range independently over the rational integers such that $|b| \leq h$. For each set of elements $X_1, \ldots, X_t$ of B, put

$$y_\mu = \sum_{\tau=1}^{t} \Omega_{\mu\tau}X_\tau, \qquad \mu = 1, \ldots, m.$$

This defines $(2h + 1)gt$ m-tuples $y_1, \ldots, y_m$, not necessarily different from one another. Also, since

$$\lceil X_\tau \rceil < \gamma h, \tag{41}$$

we have from (39) that

$$\lceil y_\mu \rceil < \gamma^m h. \tag{42}$$

Each number y_μ has a basis representation $c_1\rho_1 + \cdots + c_g\rho_g$; similar representations, with the same c_i and with the ρ_i replaced by their conjugates, hold for the conjugates of y_μ. The determinant formed from the ρ_i and their conjugates is not zero, so that it is possible to solve the g equations defining y_μ and its conjugates for the numbers c_i, giving each c_i as a linear expression in the conjugates of y_μ, with coefficients depending only on K. From (42) it follows that for $i = 1, \ldots, g$,

$$|c_i| < \gamma^m h. \tag{43}$$

There are, however, exactly $(2\gamma^m h + 1)^g$ different integers of K whose basis representation satisfies (43); therefore there are at most $(2\gamma^m h + 1)^{gm}$ different systems $y_1, \ldots, y_m$. If

$$(2\gamma^m h + 1)^{gm} < (2h + 1)^{gt}, \tag{44}$$

then two systems $y_1, \ldots, y_m$ corresponding to two different sets $X_1, \ldots, X_t$ coincide, and the respective differences $X_1 - X_1', \ldots,$ $X_t - X_t'$ constitute a solution of (33). These differences, which we call $C_1, \ldots, C_t$, are not all zero, and by (41), they satisfy the condition

$$\lceil C_\tau \rceil < \gamma h.$$

By (35), $t \geq 2m$, so (44) holds if

$$(2\gamma^m h + 1)^m < (2h + 1)^{2m},$$

which is clearly true if $h = \gamma^m$. But then (40) holds.

We now designate by m_0 a fixed value of m such that (38) holds, and by k_0 and $l_\varkappa^{(0)}$ the corresponding values of k and $l_\varkappa$, and define Φ to be a fixed function corresponding to m_0 and having all the properties described up to this point. We are now able to perform an induction.

We know that Φ possesses m_0 zeros, if each is counted with its proper multiplicity. It is asserted that *if m_0 is sufficiently large, then $\Phi(z)$ vanishes at all the points $\zeta_1, \zeta_2, \ldots$.*. This is proved inductively by showing that if $\Phi(z) = 0$ for $z = \zeta_1, \ldots, \zeta_m$, with $m \geq m_0$, then also $\Phi(\zeta_{m+1}) = 0$, if m_0 is sufficiently large. More precisely, we assume that Φ has a zero at $z_\varkappa (\varkappa = 0, \ldots, k)$ of order $l_\varkappa + 1$, with

$$\sum_{\varkappa=0}^{k} (l_\varkappa + 1) = m \geq m_0,$$

and shall deduce that Φ has a zero at $\zeta_{m+1} = \zeta$ of order $\Lambda + 1$, where

$$\Lambda = \begin{cases} 0 & \text{if } \zeta = z_{k+1} \neq z_\varkappa \text{ for } \varkappa = 0, \ldots, k, \\ l_\varkappa + 1 & \text{if } \zeta = z_\sigma \text{ and } 0 \leq \sigma \leq k. \end{cases}$$

Here $k = k(m)$ and $l_\varkappa = l_\varkappa(m)$.

Put

$$G(z) = \prod_{\nu=1}^{n} G_\nu^{l_\nu}(z);$$

then $G(z)\Phi(z)$ is an entire function which vanishes at the same points $z_\varkappa$ as $\Phi(z)$, and to the same order, by (31). We also put

$$Q(z) = \prod_{\substack{\varkappa=0 \\ z_\varkappa \neq \zeta}}^{k} (z - z_\varkappa)^{l_\varkappa+1}.$$

By Cauchy's theorem,

$$\left. \frac{d^\Lambda (G(z)\Phi(z)/Q(z))}{dz^\Lambda} \right|_{z=\zeta} = \frac{\Lambda!}{2\pi i} \int_\Gamma \frac{G(z)\Phi(z)}{Q(z)} \frac{dz}{(z - \zeta)^{\Lambda+1}}. \quad (45)$$

Here Γ is the circle

$$|z| = R_1 = R^\vartheta, \qquad \vartheta > 1,$$

where $R = r(m + 1)$ if $\alpha < \infty$ (we recall that $r(m)$ was defined earlier as the radius of the smallest circle about O containing $z_0, \ldots, z_k$), while if $\alpha = \infty$, R is so chosen that

$$R \geq r(m+1), \qquad \lim_{m \to \infty} \frac{\log m + 1}{\log R} = \infty, \qquad \lim_{m \to \infty} R = \infty.$$

Since Φ has a zero of order Λ at ζ, the left side of (45) is simply

$$\frac{G(z)}{Q(z)} \Phi^{(\lambda)}(z) \bigg|_{z=\zeta} ,$$

so that

$$|\Phi^{(\lambda)}(\zeta)| = \frac{Q(\zeta)}{G(\zeta)} \frac{\Lambda!}{2\pi} \left| \int_\Gamma \frac{G(z)\Phi(z)}{Q(z)} \frac{dz}{(z-\zeta)^{\Lambda+1}} \right| . \tag{46}$$

We shall use this representation to estimate $|\Phi^{(\lambda)}(\zeta)|$. By the inequality (40), we have

$$\max_{|z|=R_1} |G(z)\Phi(z)| < t\gamma^{m_0} \exp\left(\sum_{\nu=1}^{n} t_\nu R_1^{\mu_\nu + \epsilon} \right),$$

where $\epsilon \to 0$ as $R_1 \to \infty$, or equivalently as $m \to \infty$. By the definition (28) of α,

$$R_1 = R^\vartheta < m^{\vartheta/(\alpha-\epsilon)}$$

or

$$R_1 < m^{\vartheta/\alpha+\epsilon},$$

even for $\alpha = \infty$. Hence

$$\max_{|z|=R_1} |G(z)\Phi(z)| < \gamma^{m_0} \exp\left(\sum_{\nu=1}^{n} t_\nu m^{\vartheta \eta_\nu + \epsilon} \right),$$

since it is easily seen from the definition (34) that $t < \gamma^{m_0}$. From (34) we also have that

$$\sum_{\nu=1}^{n} t_\nu m^{\vartheta \eta_\nu + \epsilon} = \sum_{\nu=1}^{n} (2^{1/n} m_0^{(1+\eta_1+\cdots+\eta_n)/n-\eta_\nu}) m^{\vartheta \eta_\nu + \epsilon}$$

$$\leq \sum_{\nu=1}^{n} 2^{1/n} m^{(1+\eta_1+\cdots+\eta_n)/n-\eta_\nu+\vartheta \eta_\nu + \epsilon}$$

$$\leq \sum_{\nu=1}^{n} 2m^{1-(\delta/n)+(\vartheta-1)\eta_\nu+\epsilon}.$$

We may suppose, with no loss in generality, that each $\eta_\nu \leq 1$. For suppose that η_n, say, is larger than 1. Then in analogy with (30),

$$\frac{\mu_1 + \cdots + \mu_{n-1}}{n-2} < \alpha,$$

and all hypotheses of the theorem are satisfied by the $n-1$ functions $f_1(z), \ldots, f_{n-1}(z)$. But if $f_1(z), \ldots, f_{n-1}(z)$ can be shown to be

algebraically dependent, then $f_1(z), \ldots, f_n(z)$ must also be dependent.

Consequently, if we put $\vartheta = 1 + \delta/2n$, we have

$$(\vartheta - 1) \max_{1 \leq \nu \leq n} (\eta_\nu) \leq \frac{\delta}{2n}.$$

If m_0, and hence also $m \geq m_0$, is so large that

$$\epsilon < \frac{\delta}{2n}, \tag{47}$$

then

$$\sum_{\nu=1}^{n} m^{1-\delta/n+(\vartheta-1)\eta_\nu+\epsilon} < nm,$$

and

$$\max_{|z|=R_1} |G(z)\Phi(z)| < \gamma^{m_0} e^{2nm} < \gamma^m. \tag{48}$$

Continuing the estimation of the right side of (46), we notice that since $\vartheta > 1$, and since R grows indefinitely with m, it is possible to choose m_0 so large that

$$\min_{|z|=R_1} |z - z_\varkappa| > \frac{R_1}{2}, \tag{49}$$

for $\varkappa = 0, \ldots, k$, and also

$$\min_{|z|=R_1} |z - \zeta| > \frac{R_1}{2}. \tag{50}$$

In that case, $$\min_{|z|=R_1} |Q(z)(z - \zeta)^{\Lambda+1}| \geq \left(\frac{R_1}{2}\right)^{m+1} \tag{51}$$

Since ζ and all the $z_\varkappa$ lie in the disk $|z| \leq R$, we have

$$|\zeta - z_\varkappa| < 2R,$$

so that $$|Q(\zeta)| < (2R)^m. \tag{52}$$

Finally, we see from (31) that

$$|G(\zeta)| > \exp\left(-\sum_{\nu=1}^{n} t_\nu m^{\eta_\nu+\epsilon}\right) > \gamma^{-m}. \tag{53}$$

Combining the relations (48), (51), (52), and (53), we have

$$|\Phi^{(\Lambda)}(\zeta)| < (2R)^m \cdot \gamma^m \Lambda^\Lambda \cdot \gamma^m \cdot \left(\frac{R_1}{2}\right)^{-m-1} \cdot R_1$$

$$< \gamma^m \Lambda^\Lambda \left(\frac{R}{R_1}\right)^m$$

$$< \gamma^m \Lambda^\Lambda R^{-(\vartheta-1)m}.$$

Since by hypothesis

$$l = \max_{0 \leq \varkappa \leq k} (l_{\varkappa}, \Lambda) \leq \frac{m+1}{\log (m+1)},$$

the inequality

$$\Lambda^{\Lambda} < \gamma^m$$

holds; hence

$$|\Phi^{(\Lambda)}(\zeta)| < \gamma^m R^{-\delta m/2n}. \tag{54}$$

Recalling that $\Phi(z)$ is a polynomial in $f_1(z), \ldots, f_n(z)$ with coefficients C_r which are integers in K, and that all the derivatives of $f_1(z), \ldots, f_n(z)$ up to order Λ have values in K for $z = \zeta$, we see that also $\Phi^{(\Lambda)}(\zeta)$ is a number in K, and that the product

$$\Phi^{(\Lambda)}(\zeta) \prod_{\nu=1}^{n} H_{\nu}^{t_{\nu}}(\zeta)$$

is an integer in K. By the same reasoning as was used in producing the estimate (36), we have

$$\prod_{\nu=1}^{n} H_{\nu}^{t_{\nu}}(\zeta) \cdot \overline{|\Phi^{(\Lambda)}(\zeta)|} < \gamma^{m_0} \cdot \prod_{\nu=1}^{n} H_{\nu}^{t_{\nu}}(\zeta) \cdot \left(\sum_{\nu=1}^{n} t_{\nu} \right)^{\Lambda}$$

$$\cdot \prod_{\nu=1}^{n} \exp (t_{\nu} m^{\eta_{\nu}+\epsilon}) \cdot \prod_{\nu=1}^{n} (t_{\nu} + 1).$$

The factor γ^{m_0} comes from the estimate (40) for $\overline{|C_r|}$, while the last product is the total number of terms in $\Phi(z)$ itself, which was an unnecessary factor in (36), where we were estimating the terms in a derivative arising from a single term in $\Phi(z)$. By arguments used previously, it follows from the last inequality that

$$\prod_{\nu=1}^{n} H_{\nu}^{t_{\nu}}(\zeta) \overline{|\Phi^{(\Lambda)}(\zeta)|} < \gamma^m.$$

Combining this with (54), we have

$$\left| \mathbf{N} \left(\Phi^{(\Lambda)}(\zeta) \prod_{\nu=1}^{n} H_{\nu}^{t_{\nu}}(\zeta) \right) \right| < \gamma^{gm} R^{-\delta m/2n}, \tag{55}$$

and the upper bound here is smaller than 1 for m sufficiently large, say $m \geq m_1$. Hence if m_0 is so large that $m_0 > m_1$, and the inequalities (47), (49), and (50) hold, then it follows from (55) that

$$\Phi^{(\Lambda)}(\zeta) = 0,$$

as asserted.

To complete the proof of Theorem 5–9, we shall make use of the following general considerations. Let $\varphi(z)$ be analytic in the disk bounded by a circle Γ, and let $x_1, \ldots, x_p$ be interior points of this disk. Then $\varphi(z)$ has an expansion

$$\varphi(z) = a_0 + a_1(z - x_1) + a_2(z - x_1)(z - x_2) + \cdots$$
$$+ a_{p-1}(z - x_1) \cdots (z - x_{p-1}) + (z - x_1) \cdots (z - x_p)R_p(z) \quad (56)$$

with constants $a_0, \ldots, a_{p-1}$ and a function $R_p(z)$ regular in the disk. In fact, if we put

$$\frac{1}{2\pi i} \int_\Gamma \frac{\varphi(t)}{(t - x_1) \cdots (t - x_{q+1})} \, dt = a_q, \quad (57)$$

for $q = 0, \ldots, p - 1$, and

$$\frac{1}{2\pi i} \int_\Gamma \frac{\varphi(t)}{(t - z)(t - x_1) \cdots (t - x_p)} \, dt = R_p(z), \quad (58)$$

then

$$\varphi(z) - \left(a_0 + a_1(z - x_1) + \cdots + a_{p-1}(z - x_1) \cdots (z - x_{p-1})\right)$$

$$= \frac{1}{2\pi i} \int_\Gamma \left(\frac{1}{t - z} - \frac{1}{t - x_1} - \frac{z - x_1}{(t - x_1)(t - x_2)} - \cdots\right.$$

$$\left. - \frac{(z - x_1) \cdots (z - x_{p-1})}{(t - x_1) \cdots (t - x_p)}\right) \varphi(t)dt$$

$$= \frac{1}{2\pi i} \int_\Gamma \frac{(z - x_1) \cdots (z - x_p)}{(t - z)(t - x_1) \cdots (t - x_p)} \varphi(t)dt$$

$$= (z - x_1) \cdots (z - x_p)R_p(z),$$

and it is clear from (58) that $R_p(z)$ is regular inside Γ.

We apply this with $\varphi(z) = \Phi(z)G(z)$, the "interpolation points" $x_1, \ldots, x_p$ being $\zeta_1, \ldots, \zeta_m$, in this order. For Γ we choose the circle $|z| = R_1 = R^\vartheta$ with $\vartheta > 1$, where $R \geq r(m)$, $\lim\limits_{m \to \infty} R = \infty$, and

$$\liminf_{m \to \infty} \frac{\log m}{\log R} = \alpha.$$

Since $\Phi(z)$ vanishes at all the points $\zeta_1, \zeta_2, \ldots$, the integrand in the expression (57) for a_q is regular in Γ, so that

$$a_q = 0 \quad \text{for } q = 0, \ldots, m - 1.$$

Hence for fixed z with $z \neq \zeta_1, \zeta_2, \ldots,$

$$G(z)\Phi(z) = \lim_{m \to \infty} \left(Q_m(z) \cdot \frac{1}{2\pi i} \int_\Gamma \frac{G(t)\Phi(t)}{Q_m(t)} \frac{dt}{t - z} \right),$$

where

$$Q_m(z) = \prod_{\varkappa=0}^{k} (z - z_\varkappa)^{l_\varkappa + 1}, \qquad \sum_{\varkappa=0}^{k} (l_\varkappa + 1) = m.$$

As in the derivation of (54), we have

$$|G(z)\Phi(z)| \leq \gamma^{m_0} \prod_{\nu=1}^{n} (t_\nu + 1) \exp\left(\sum_{\nu=1}^{n} t_\nu R_1^{\mu_\nu + \epsilon} \right) (2r)^m \left(\frac{R_1}{2} \right)^{-m}$$

$$< \gamma^m \left(\frac{R}{R_1} \right)^m = \gamma^m R^{-(\vartheta - 1)m}.$$

Since this inequality holds for arbitrarily large m, and since R increases indefinitely with m while γ does not, it must be that

$$G(z)\Phi(z) = 0.$$

Hence $\Phi(z)$ vanishes for all z, which is the assertion of the theorem.

5–7 The Hilbert-Gelfond-Schneider theorem. As an application of Schneider's theorem of the preceding section, we now prove

THEOREM 5–10. *If a and b are algebraic numbers, b is irrational, and a is neither 0 nor 1, then a^b is transcendental.*

This theorem settles a question raised by Euler concerning the arithmetic nature of the logarithm of a rational number to a rational base, and repeated in more general form in the seventh of Hilbert's famous list of 23 outstanding problems which seemed to him to be both difficult and important. The list appeared in 1900, but it was not until 1929 that Gelfond made the first contribution to the solution of this problem. Further partial results were obtained by Kusmin, Siegel, and Boehle, and in 1934 complete proofs were given almost simultaneously by Gelfond and Schneider. As mentioned earlier, the proof to be given now is most nearly in the spirit of Gelfond's 1929 paper; it should be instructive to the reader also to examine the original complete proofs by Gelfond and Schneider.

We apply Theorem 5–9 with $n = 2, f_1(z) = a^z, f_2(z) = z,$ and $\zeta_m = u + vb$, where u and v range over the positive integers. On

account of the irrationality of b, the numbers ζ_m are distinct; they are to be ordered by the size of $u + v$, and otherwise arbitrarily. Suppose that in the sequence $\zeta_1, \ldots, \zeta_m$, all the numbers occur for which $u + v \leq d$, and possibly some (but not all) of those for which $u + v = d + 1$. Then clearly

$$\frac{d(d-1)}{2} \leq m < \frac{d(d+1)}{2},$$

while

$$r = \max(|u + vb|) \leq (d+1)(1 + |b|),$$

and (taking $u = d - 1$, $v = \pm 1$),

$$r \geq d - 1 - |b|.$$

These inequalities show that $\gamma_1 r^2 < m < \gamma_2 r^2$, for some positive γ_1 and γ_2. Thus

$$\alpha = \lim_{m \to \infty} \frac{\log m}{\log r} = 2.$$

By the choice of $f_1(z)$ and $f_2(z)$, $\mu_1 = 1$ and $\mu_2 = 0$, and the inequality (30) holds. Since a^z and z are entire, (31) is without force. If we suppose that z and a^z are elements of an algebraic number field K for $z = b$, then $f_2(\zeta) = u + vb$ and $f_1(\zeta) = a^u(a^b)^v$ are also in K for positive integral u and v. (We need not examine the derivatives, since $\zeta_1, \zeta_2, \ldots$ are distinct.) Moreover, if c is a positive rational integer such that ca, cb, and ca^b are integers of K, then we can choose

$$H_1(z_\kappa) = c^{u+v} \quad \text{and} \quad H_2(z_\kappa) = c$$

for $z_\kappa = u + vb$. It follows from this and the definitions of $f_1(z)$ and $f_2(z)$ that the inequality (32) holds. Thus, under the assumption that a, b, and a^b are all algebraic, all hypotheses of Theorem 5-9 are satisfied, and it follows that z and a^z are algebraically dependent. This being palpably false, the above assumption cannot be maintained, and the theorem is proved.

PROBLEM

Show that e^ϑ is transcendental for algebraic $\vartheta \neq 0$. [*Hint:* Choose $f_1(z) = e^z$, $f_2(z) = z$, and

$$\zeta_{(n-1)^2+1} = \cdots = \zeta_{n^2} = n\vartheta,$$

for $n = 1, 2, \ldots$.]

REFERENCES

Section 5-1

I. Niven, *Bulletin of the American Mathematical Society* **53**, 509 (1947).

Section 5-2

J. Liouville, *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* (Paris) **18**, 883–885, 910–911 (1844); *Journal des Mathematiques Pures et Appliquées* (Paris) **16**, 133–142 (1851).

Sections 5-3, 5-4, 5-5

Most of this material is adapted from Mahler, *Journal für die Reine und Angewandte Mathematik* (Berlin) **66**, 117–150 (1932). For the existence of U-numbers of each degree, see LeVeque, *Journal of the London Mathematical Society* **28**, 220–229 (1953).

Section 5-6

Siegel's work on Bessel functions is to be found in *Abhandlung der Kgl. Preussischen Akademie der Wissenschaften* (Berlin), article no. 1, 70 pp. (1929). The first result stated concerning the $\wp$-function is due to Siegel, *Journal für die Reine und Angewandte Mathematik* **167**, 62–69 (1932); the second to Schneider, *ibid.*, **172**, 70–74 (1934). The transcendence of decimals formed from polynomial values was proved by Mahler, *Proceedings Konink. Nederlandsche Akademie van Wetenschappen* (Amsterdam) **40**, 421–428 (1937), and that of the series $\sum [n\omega]z^n$ by Mahler, *Mathematische Annalen* (Leipzig) **101**, 342–366 (1929) and **103**, 532 (1930), and *Mathematische Zeitschrift* (Berlin) **32**, 545–585 (1930).

Schneider's Theorem 5-9 appeared in *Mathematische Annalen* **121**, 131–140 (1949); he includes a bibliography of work on integral-valued functions. Pólya's work appeared in *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, pp. 1–10 (1920), and Gelfond's in *Tôhoku Mathematical Journal* (Sendai, Japan) **30**, 280–285 (1929).

Section 5-7

Hilbert's problems appeared in *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, pp. 253–297 (1900). The complete solution of the seventh was given by Gelfond, *Comptes Rendus de l'Académie des Sciences de l' U.R.S.S.* (Moscow) **2**, 1–3 (in Russian), 4–6 (in French) (1934), and *Bulletin de l'Académie des Sciences de l' U.R.S.S.* (Leningrad) **7**, 623–640 (1934); and by Schneider, *Journal für die Reine und Angewandte Mathematik* **172**, 65–69 (1934). There is an excellent exposition of Gelfond's method by E. Hille, *American Mathematical Monthly* **49**, 654–661 (1942).

CHAPTER 6

DIRICHLET'S THEOREM

In this chapter and the next we shall consider various questions concerning the distribution of the rational primes. This is a large and difficult field, and we shall be able to obtain only a few of the important results. The first of them, to which this chapter is devoted, is Dirichlet's famous theorem that there are infinitely many primes of the form $km + l$, where k and l are fixed integers which are relatively prime.

6-1 Introduction. Although proofs of certain special cases of Dirichlet's theorem are given in elementary texts,* the methods used cannot be generalized to prove the full theorem. To get an idea of the method used by Dirichlet, let us consider the question of the infinitude of the set of primes of the form $4k + 1$. We base the discussion on the *Riemann ζ-function*, defined for $s > 1$ by the equation

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

This is perhaps the simplest of all the *Dirichlet series*

$$\sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

which play an important role in prime number theory. One reason for their importance is exhibited in the following theorem, which gives a relation between the set of primes and the set of positive integers.

THEOREM 6-1. For $s > 1$,

$$\zeta(s) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}. \tag{1}$$

Proof: In less abbreviated form, the assertion is that

$$\lim_{N \to \infty} \prod_{p \le N} \left(1 - \frac{1}{p^s}\right)^{-1} = \lim_{N \to \infty} \sum_{n=1}^{N} \frac{1}{n^s}.$$

---

* See, for example, Volume I, pp. 9, 46, 59.

The relation

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{n=0}^{\infty} x^n$$

holds for $|x| < 1$; since $|p^{-s}| < 1$, we have

$$\prod_{p \leq N} (1 - p^{-s})^{-1} = \prod_{p \leq N} (1 + p^{-s} + p^{-2s} + \cdots).$$

Multiplying out the product on the right, we obtain terms of the form n^{-s}, where n runs over the integers composed exclusively of primes not exceeding N. Moreover, each such n occurs exactly once, by the Unique Factorization Theorem. The multiplication is permissible, since the series involved are absolutely convergent, and the terms can be arranged in any order. Thus

$$\prod_{p \leq N} (1 - p^{-s})^{-1} = \sum' n^{-s},$$

where the accent indicates a summation, in the natural order, over all n such that $p|n$ implies $p \leq N$. In particular, the sum contains all terms n^{-s} for which $n \leq N$. Hence

$$\prod_{p \leq N} (1 - p^{-s})^{-1} = \sum_{n=1}^{N} n^{-s} + \sum_{n>N}' n^{-s},$$

and

$$0 < \sum_{n>N}' n^{-s} \leq \sum_{n=N+1}^{\infty} n^{-s} < \int_{N}^{\infty} x^{-s}\, dx = \frac{1}{(s-1)N^{s-1}},$$

since $s > 1$. Thus

$$\sum_{n>N}' n^{-s} = o(1)$$

as $N \to \infty$, and

$$\lim_{N\to\infty} \prod_{p \leq N} (1 - p^{-s})^{-1} = \lim_{N\to\infty} \sum_{n=1}^{N} n^{-s} = \zeta(s).$$

To see exactly how $\zeta(s)$ behaves as $s \to 1^+$, we use the following standard result.*

LEMMA. *Suppose that $\lambda_1, \lambda_2, \ldots$ is a nondecreasing sequence tending to infinity, that $c_1, c_2, \ldots$ is an arbitrary sequence of real or complex numbers, and that $f(x)$ has a continuous derivative for $x \geq \lambda_1$. Put*

$$C(x) = \sum_{\substack{n \\ \lambda_n \leq x}} c_n.$$

---

* See, for example, Volume I, Theorem 6–15.

Then for $x \geq \lambda_1$,

$$\sum_{\lambda_n \leq x} c_n f(\lambda_n) = C(x)f(x) - \int_{\lambda_1}^{x} C(t)f'(t) \, dt.$$

Applying this with $\lambda_n = n, c_n = 1, f(x) = x^{-s}$, we obtain

$$\sum_{n \leq x} \frac{1}{n^s} = s \int_1^x \frac{[t]}{t^{s+1}} \, dt + \frac{[x]}{x^s}$$

for $x \geq 1$. If we put $(x) = x - [x]$, we have for $s > 1$,

$$\sum_{n \leq x} \frac{1}{n^s} = s \int_1^x \frac{t - (t)}{t^{s+1}} \, dt + \frac{x - (x)}{x^s}$$

$$= s \int_1^x \frac{dt}{t^s} - s \int_1^x \frac{(t)}{t^{s+1}} \, dt + \frac{1}{x^{s-1}} - \frac{(x)}{x^s}$$

$$= \frac{s}{s-1} - \frac{s}{(s-1)x^{s-1}} - s \int_1^x \frac{(t)}{t^{s+1}} \, dt + \frac{1}{x^{s-1}} - \frac{(x)}{x^s} \cdot$$

Letting x increase without bound and noting that $0 \leq (x) < 1$, we have

$$\zeta(s) = \frac{s}{s-1} - s \int_1^\infty \frac{(t)}{t^{s+1}} \, dt. \qquad (2)$$

This expression for $\zeta(s)$ agrees with the earlier definition for $s > 1$, but it is also meaningful for $0 < s < 1$, since the integral converges for all $s > 0$. It may therefore be thought of as defining $\zeta(s)$ for $s > 0, s \neq 1$. At any rate, (2) shows that

$$\lim_{s \to 1^+} \zeta(s)(s-1) = 1, \qquad (3)$$

and *a fortiori* that

$$\lim_{s \to 1^+} \zeta(s) = \infty. \qquad (4)$$

For the remainder of this section, let q and r designate primes of the forms $4k + 1$ and $4k - 1$ respectively. Define the function $\chi(n)$ by the equations

$$\chi(1) = 1, \qquad \chi(q) = 1, \qquad \chi(r) = -1, \qquad \chi(2) = 0,$$

$$\chi(mn) = \chi(m)\chi(n) \text{ for every pair of integers } m, n.$$

(A function which satisfies the last of these conditions is said to be *completely multiplicative;* it is entirely determined when its values for all prime arguments are known, since $\chi(p^\alpha) = (\chi(p))^\alpha$.) Inasmuch

as $n \equiv 1 \pmod 4$ if and only if $2 \nmid n$ and the total number of r's dividing n is even, we have

$$\chi(n) = \begin{cases} 0 & \text{if } 2|n, \\ (-1)^{\frac{1}{2}(n-1)} & \text{if } 2 \nmid n. \end{cases}$$

We now investigate the function

$$L(s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

If we write $\sum a_n \ll \sum b_n$ to mean that $|a_n| \leq b_n$ for $n = 1, 2, \ldots$, then

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \ll \sum_{n=1}^{\infty} \frac{1}{n^s}$$

for $s > 1$, so that the series for $L(s)$ is absolutely convergent for $s > 1$. More than this is true, however: *the series for $L(s)$ converges for $s > 0$.* For we note that for any $n > 0$,

$$\chi(n) + \chi(n+1) + \chi(n+2) + \chi(n+3) = 0,$$

so that we have

$$\sum_{n=1}^{N} \chi(n)$$

$$= \sum_{n=1}^{4} \chi(n) + \sum_{n=5}^{8} \chi(n) + \cdots + \sum_{n=4[\frac{1}{4}N]-3}^{4[\frac{1}{4}N]} \chi(n) + \sum_{n=4[\frac{1}{4}N]+1}^{N} \chi(n)$$

$$= 0 + 0 + \cdots + 0 + \sum_{n=4[\frac{1}{4}N]+1}^{N} \chi(n),$$

and hence

$$\left| \sum_{n=1}^{N} \chi(n) \right| \leq 1.$$

The truth of the assertion is therefore a weak consequence of the following theorem, which is due to Abel.

THEOREM 6–2. *If $\{a_n\}$ is a sequence of constants for which*

$$\sum_{n=1}^{N} a_n = O(1)$$

as $N \to \infty$, and if $\{b_n(s)\}$ is a sequence of positive-valued functions which converges monotonically and uniformly to zero for s in some

interval J, then the series

$$\sum_{n=1}^{\infty} a_n b_n(s)$$

converges uniformly for s in J.

Proof: Put

$$A_n = \sum_{k=1}^{n} a_k,$$

so that $|A_n| < A$ for some A and all n. Using the monotonicity of $b_n(s)$, we have

$$\left| \sum_{n=j}^{k} a_n b_n(s) \right| = \left| \sum_{n=j}^{k} (A_n - A_{n-1}) b_n(s) \right|$$

$$= \left| \sum_{n=j}^{k-1} A_n(b_n(s) - b_{n+1}(s)) + A_k b_k(s) - A_{j-1} b_j(s) \right|$$

$$\leq A(b_j(s) - b_k(s)) + A b_k(s) + A b_j(s) = 2A b_j(s).$$

By hypothesis, this upper bound can be made uniformly small, for s in J, by taking j sufficiently large. This proves the theorem.

Here we have a situation which does not arise in the case of power series. For while a power series converges absolutely at every interior point of its interval of convergence, the Dirichlet series for $L(s)$ converges for $s > 0$, but converges absolutely only for $s > 1$, since the series

$$\sum_{n=1}^{\infty} \left| \frac{\chi(n)}{n} \right| = \sum_{n=0}^{\infty} \frac{1}{2n+1}$$

diverges.

On account of the complete multiplicativity of χ, we have

$$\left(1 - \frac{\chi(p)}{p^s} \right)^{-1} = 1 + \frac{\chi(p)}{p^s} + \frac{(\chi(p))^2}{p^{2s}} + \cdots$$

$$= 1 + \frac{\chi(p)}{p^s} + \frac{\chi(p^2)}{p^{2s}} + \cdots.$$

Using this idea, the proof of Theorem 6–1 can easily be modified to yield

THEOREM 6–3. *If f is completely multiplicative, and the series*

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

converges absolutely for s > s₀, then

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \prod_{p} \left(1 - \frac{f(p)}{p^s}\right)^{-1}$$

for s > s₀.

COROLLARY. *For s > 1,*

$$L(s) = \prod_{p} \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}.$$

We are finally in a position to prove Dirichlet's theorem for primes of the form $4k + 1$. Let s be greater than 1. We have

$$\zeta(s) = \prod_{p} (1 - p^{-s})^{-1} = (1 - 2^{-s})^{-1} \prod_{q} (1 - q^{-s})^{-1} \prod_{r} (1 - r^{-s})^{-1},$$

and, from the corollary to Theorem 6–3,

$$L(s) = \prod_{q} (1 - q^{-s})^{-1} \prod_{r} (1 + r^{-s})^{-1}.$$

Hence

$$\zeta(s)L(s) = (1 - 2^{-s})^{-1} \prod_{q} (1 - q^{-s})^{-2} \prod_{r} (1 - r^{-2s})^{-1}. \quad (5)$$

Now, for $s \geq 1$,

$$L(s) = 1 - \frac{1}{3^s} + \frac{1}{5^s} - \frac{1}{7^s} + \cdots > 1 - \frac{1}{3^s} \geq \frac{2}{3},$$

and so

$$\lim_{s \to 1^+} \zeta(s)L(s) = \infty,$$

by (4). If there were only finitely many primes q, the expression on the right side of (6) would remain bounded as $s \to 1^+$, since for $s \geq 1$,

$$\prod_{r} (1 - r^{-2s})^{-1} \leq \prod_{r} (1 - r^{-2})^{-1} \leq \prod_{p} (1 - p^{-2})^{-1} = \zeta(2).$$

This contradiction shows not only that there are infinitely many primes q, but also that they occur sufficiently frequently that

$$\lim_{s \to 1^+} \prod_{q} (1 - q^{-s})^{-1} = \infty.$$

The proof which has just been given contains most of the essential features of the general proof. The major formal difference which will arise in the general case is that we shall have to consider a number of functions like χ above, and each will have an associated Dirichlet

series, some aspects of whose behavior must be investigated. The most difficult part of the proof lies in showing that these series do not vanish at $s = 1$, a point which caused no trouble in the present case.

Let $b_n = 1$ or 0, according as the equation $n = x^2 + y^2$ has or does not have a solution in integers x, y. It is known* that $b_n = 1$ if and only if every prime $r \equiv -1 \pmod 4$ which divides n occurs to an even power in the canonical factorization of n. Show that the series

$$\sum_{n=1}^{\infty} \frac{b_n}{n^s}$$

converges for $s > 1$, and diverges for $s \leq 1$. [*Hint:* Establish a relation among $\zeta(s)$, $L(s)$, and the square of the given series.]

6–2 Characters. We recall that the elements of a reduced residue system (mod k) form an abelian group under multiplication (mod k), which we designate by $M(k)$. The number of elements of $M(k)$, called its *order*, is $\varphi(k)$; hereafter we shall use h as an abbreviation for $\varphi(k)$.

One of the fundamental theorems on finite abelian groups is that every such group has a *basis:* if it is a multiplicative group, this means that there is a set of elements $A_1, \ldots, A_r$ such that every element of the group can be written uniquely in the form

$$A_1^{x_1} \cdots A_r^{x_r},$$

where each x_i is one of the integers $0, 1, \ldots, \operatorname{ord} A_i - 1$, and $\operatorname{ord} A_i$ is the order of the cyclic subgroup generated by A_i. Moreover, the product of all the numbers $\operatorname{ord} A_i$ is the order of the group. The following theorem, for which we give a proof based on the theory of primitive roots† is a special case.

THEOREM 6–4.   (a) *Let* $k = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, *where* $p_i \neq p_j$ *and each of the prime powers* $p_i^{\alpha_i}$ *has a primitive root, say* g_i. *Then the numbers*

---

* See, for example, Volume I, Theorem 7–3.

† See, for example, Volume I, Chapter 4.

$A_1, \ldots, A_r$ *form a basis for* $M(k)$ *if, for each* i,

$$A_i \equiv \begin{cases} g_i \pmod{p_i^{\alpha_i}} \\ 1 \pmod{p_j^{\alpha_j}} \end{cases} \quad if \ j \neq i, \ \ 1 \leq j \leq r.$$

(b) *Let* $k = 2^{\alpha} p_2^{\alpha_2} \cdots p_r^{\alpha_r}$, *where* $\alpha \geq 3$, *and let* g_i *be a primitive root of* $p_i^{\alpha_i}$ *for* $2 \leq i \leq r$. *Then the numbers* $A_0, A_1, \ldots, A_r$ *constitute a basis for* $M(k)$, *where*

$$A_0 \equiv \begin{cases} -1 \pmod{2^{\alpha}} \\ 1 \pmod{p_i^{\alpha_i}} \end{cases} \quad for \ 2 \leq i \leq r,$$

$$A_1 \equiv \begin{cases} 5 \pmod{2^{\alpha}} \\ 1 \pmod{p_i^{\alpha_i}} \end{cases} \quad for \ 2 \leq i \leq r,$$

and for $2 \leq i \leq r$,

$$A_i \equiv \begin{cases} g_i \pmod{p_i^{\alpha_i}} \\ 1 \pmod{p_j^{\alpha_j}} \end{cases} \quad for \ j \neq i, \ \ 1 \leq j \leq r.$$

Proof: Let a be relatively prime to k. Then it is also prime to every divisor of k, so that there are unique elements $a_1, \ldots, a_r$ of $M(p_1^{\alpha_1}), \ldots, M(p_r^{\alpha_r})$, respectively, such that

$$\begin{aligned} a &\equiv a_1 \pmod{p_1^{\alpha_1}}, \\ &\ \ \vdots \\ a &\equiv a_r \pmod{p_r^{\alpha_r}}. \end{aligned} \tag{6}$$

Conversely, for any choice of $a_1, \ldots, a_r$ in $M(p_1^{\alpha_1}), \ldots, M(p_r^{\alpha_r})$, respectively, the system (6) has a solution a which is unique modulo k, by the Chinese Remainder Theorem, and a is prime to k. Moreover, if a is the solution of (6), and if, for $1 \leq i \leq r$, b_i is the solution of the system

$$b_i \equiv \begin{cases} a_i \pmod{p_i^{\alpha_i}} \\ 1 \pmod{p_j^{\alpha_j}} \end{cases} \quad for \ j \neq i, \ \ 1 \leq j \leq r, \tag{7}$$

then

$$b_1 \cdots b_r \equiv 1 \cdots 1 \cdot a_i \cdot 1 \cdots 1 \equiv a_i \pmod{p_i^{\alpha_i}}, \quad for \ 1 \leq i \leq r,$$

so that

$$a \equiv b_1 \cdots b_r \pmod{k}. \tag{8}$$

(Thus, in the language of group theory, $M(k)$ is the direct product of $M(p_1^{\alpha_1}), \ldots, M(p_r^{\alpha_r})$.)

Now if $p_i^{\alpha_i}$ has a primitive root g_i, and

$$A_i \equiv \begin{cases} g_i \ (\mathrm{mod}\ p_i^{\alpha_i}) \\ 1 \ (\mathrm{mod}\ p_j^{\alpha_j}) \end{cases} \qquad \text{for } j \neq i, \quad 1 \leq j \leq r,$$

then, since

$$a_i \equiv g_i^{\,\mathrm{ind}\, a} \ (\mathrm{mod}\ p_i^{\alpha_i}),$$

we have that

$$b_i \equiv A_i^{\,\mathrm{ind}\, a_i} \ (\mathrm{mod}\ p_j^{\alpha_j}), \qquad \text{for } 1 \leq j \leq r,$$

and hence

$$b_i \equiv A_i^{\,\mathrm{ind}\, a_i} \ (\mathrm{mod}\ k).$$

Thus by the congruence (8), if all $p_i^{\alpha_i}$ have primitive roots,

$$a \equiv A_1^{\,\mathrm{ind}\, a_1} \cdots A_r^{\,\mathrm{ind}\, a_r} \ (\mathrm{mod}\ k),$$

and this representation is unique if each index is given its smallest non-negative value, so that $0 \leq \mathrm{ind}\, a_i < \varphi(p_i^{\alpha_i})$.

On the other hand, if $p_1^{\alpha_1} = 2^\alpha$ with $\alpha \geq 3$, then -1 and 5 constitute a basis for $M(p_1^{\alpha_1})$. For* 5 is a primitive λ-root of 2^α, so that the $2^{\alpha-2}$ numbers

$$5, 5^2, \ldots, 5^{2^{\alpha-2}}$$

are distinct $(\mathrm{mod}\ 2^\alpha)$; since they are all congruent to $1 \ (\mathrm{mod}\ 4)$, and since there are exactly $2^{\alpha-2}$ numbers in a reduced residue system $(\mathrm{mod}\ 2^\alpha)$ which are congruent to $1 \ (\mathrm{mod}\ 4)$, these must be the numbers. Likewise, their negatives are all the numbers congruent to $-1 \ (\mathrm{mod}\ 4)$ in a reduced residue system $(\mathrm{mod}\ 2^\alpha)$.† Hence, if a is in $M(2^\alpha)$, then, for some choice of x_0 and x_1,

$$a \equiv (-1)^{x_0} 5^{x_1} \ (\mathrm{mod}\ 2^\alpha).$$

Thus if $A_0, \ldots, A_r$ are defined as in part (b) of the theorem, we have

$$a \equiv A_0^{x_0} A_1^{x_1} A_2^{\,\mathrm{ind}\, a_2} \cdots A_r^{\,\mathrm{ind}\, a_r} \ (\mathrm{mod}\ k),$$

and the representation is again unique if we require that

$$0 \leq x_0 < \mathrm{ord}\, A_0 = 2,$$

$$0 \leq x_1 < \mathrm{ord}\, A_1 = 2^{\alpha-2},$$

$$0 \leq \mathrm{ind}\, a_i < \mathrm{ord}\, A_i = \varphi(p_i^{\alpha_i}).$$

---

* See Volume I, Theorem 4–9.

† A similar argument is used in Volume I, in the proof of Theorem 5–1.

Notice that in the two cases we have

$$\operatorname{ord} A_1 \cdots \operatorname{ord} A_r = \varphi(p_1{}^{\alpha_1}) \cdots \varphi(p_r{}^{\alpha_r}) = h,$$

$$\operatorname{ord} A_0 \cdot \operatorname{ord} A_1 \cdots \operatorname{ord} A_r = 2 \cdot 2^{\alpha-2} \cdot \varphi(p_2{}^{\alpha_2}) \cdots \varphi(p_r{}^{\alpha_r}) = h.$$

To obviate the distinction between cases (a) and (b), we rename the basis elements $B_1, \ldots, B_m$, and put $\operatorname{ord} B_i = h_i$ for $i = 1, \ldots, m$.

A complex-valued function χ, defined over the group $M(k)$ (more generally, over any finite abelian group), is called a *character* (mod k) (or a *character of the group*) if it is completely multiplicative and not identically zero, that is, if

$$\chi(ab) = \chi(a)\chi(b), \qquad \text{for } a \text{ and } b \text{ in } M(k),$$

$$\chi(a) \neq 0, \qquad \text{for some } a \text{ in } M(k).$$

Since in the group $M(k)$ we identify integers which are congruent (mod k), we have

$$\chi(a) = \chi(a'), \qquad \text{if } a \equiv a' \pmod{k} \quad \text{and} \quad (a, k) = (a', k) = 1,$$

so that one could also think of characters as being defined over the residue classes themselves. Notice that necessarily $\chi(1) = 1$, since for any a for which $\chi(a) \neq 0$, we have $\chi(a) = \chi(a \cdot 1) = \chi(a)\chi(1)$. Moreover, if a is in $M(k)$ and $\operatorname{ord} a = t$, then

$$(\chi(a))^t = \chi(a^t) = \chi(1) = 1.$$

Since $t|h$, it follows that every value of every character is an hth root of unity.

On account of its complete multiplicativity, any character is totally determined when its value is specified for each basis element B_j. Thus the characters are contained in the set of all completely multiplicative functions over $M(k)$ for which

$$\chi(B_j) = e^{2\pi i \beta_j / h_j}, \qquad 0 \le \beta_j < h_j, \tag{9}$$

for $j = 1, \ldots, m$. But conversely, every such function is obviously a character, and different choices of the β's lead to different characters. Thus there are h different characters, corresponding to the $h_1 \cdots h_m$ different m-tuples $(\beta_1, \ldots, \beta_m)$.

Two groups G and G', with elements $a, b, \ldots$ and $a', b', \ldots$, are said to be *isomorphic* if it is possible to find a pairing of elements of G with elements of G', such that each element of G corresponds to

precisely one element of G', and conversely, and such that if $a \leftrightarrow a'$ and $b \leftrightarrow b'$, then $ab \leftrightarrow a'b'$. In this case the groups are abstractly identical, and any theorem concerning one group has an immediate analog for the other group. To construct such an isomorphism between two finite abelian groups, it suffices to find a one-to-one correspondence of basis elements such that corresponding elements have equal orders. For let the bases be $C_1, \ldots, C_s$ and $C_1', \ldots, C_s'$, so named that ord $C_i = $ ord C_i', for $i = 1, \ldots, s$. Then we can make a and a' correspond if

$$a = C_1^{x_1} \cdots C_s^{x_s} \quad \text{and} \quad a' = C_1'^{x_1} \cdots C_s'^{x_s},$$

$$0 \leq x_i < \text{ord } C_i.$$

For if also

$$b = C_1^{y_1} \cdots C_s^{y_s} \quad \text{and} \quad b' = C_1'^{y_1} \cdots C_s'^{y_s},$$

then

$$ab = C_1^{x_1+y_1} \cdots C_s^{x_s+y_s}, \quad a'b' = C_1'^{x_1+y_1} \cdots C_s'^{x_s+y_s},$$

and

$$(ab)' = C_1'^{x_1+y_1} \cdots C_s'^{x_s+y_s} = a'b'.$$

Moreover, this is a one-to-one correspondence, since the representations by basis elements are unique for the ranges $0 \leq x_i < $ ord $C_i = $ ord C_i', $1 \leq i \leq s$.

For the basis $B_1, \ldots, B_m$ of $M(k)$, define characters $\chi_1, \ldots, \chi_m$ as follows:

$$\chi_\mu(B_j) = \begin{cases} e^{2\pi i/h_\mu} & \text{if } j = \mu, \\ 1 & \text{if } j \neq \mu, \quad 1 \leq j \leq m. \end{cases} \tag{10}$$

Then from the sentence containing equation (9), we see that every character can be represented uniquely in the form

$$\chi = \chi_1^{\beta_1} \cdots \chi_m^{\beta_m}. \quad 0 \leq \beta_i < h_i \quad \text{for } i = 1, \ldots, m,$$

since this gives $\chi(B_j)$ as in (9). (We say that two characters are equal if they have the same value for every element of the group, and define the product of two characters as the function whose values are the products of the component values; this function is also a character by the sentence following (9).) Under multiplication, the characters form a group $X(k)$, having basis $\chi_1, \ldots, \chi_m$; since ord $\chi_i = $ ord B_i, the groups $X(k)$ and $M(k)$ are isomorphic. The

unit element of $X(k)$ is the character χ_0, the *principal character*, such that $\chi_0(a) = 1$ for every a in $M(k)$.

We summarize the chief results obtained so far.

THEOREM 6–5. *There are h distinct characters (mod k), and these form a group $X(k)$ which is isomorphic to $M(k)$. Every value $\chi(a)$ is an hth root of unity. The characters $\chi_1, \ldots, \chi_m$ defined in (10) form a basis for $X(k)$.*

We shall also need the following result.

THEOREM 6–6. *If χ is in $X(k)$, then*

$$\sum_{a \in M(k)} \chi(a) = \begin{cases} h & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \neq \chi_0, \end{cases}$$

while if a is in $M(k)$, then

$$\sum_{\chi \in X(k)} \chi(a) = \begin{cases} h & \text{if } a \equiv 1 \ (\text{mod } k), \\ 0 & \text{if } a \not\equiv 1 \ (\text{mod } k). \end{cases}$$

Proof: We have

$$\sum_{a \in M(k)} \chi_0(a) = \sum_{a \in M(k)} 1 = h.$$

If $\chi \neq \chi_0$, then for some $\bar{a}$ in $M(k)$, $\chi(\bar{a}) \neq 1$. For this $\bar{a}$,

$$\chi(\bar{a}) \sum_a \chi(a) = \sum_a \chi(a)\chi(\bar{a}) = \sum_a \chi(a\bar{a}),$$

and, as a runs over a reduced residue system, so does $a\bar{a}$, so that

$$\chi(\bar{a}) \sum_a \chi(a) = \sum_a \chi(a),$$

$$\sum_a \chi(a) = 0.$$

If $a \not\equiv 1 \ (\text{mod } k)$, and $a = B_1^{x_1} \cdots B_m^{x_m}$, then some $x_i \neq 0$. For this i, $\chi_i(a) \neq 1$, and

$$\chi_i(a) \sum_\chi \chi(a) = \sum_\chi \chi_i(a)\chi(a) = \sum_\chi \chi_i'(a),$$

where $\chi_i' = \chi_i\chi$. As χ runs over $X(k)$, so also does $\chi_i\chi = \chi_i'$, and

$$\chi_i(a) \sum_\chi \chi(a) = \sum_\chi \chi(a),$$

$$\sum_\chi \chi(a) = 0.$$

$\chi(a)$ has so far been defined only for arguments relatively prime to k. For simplicity in later formulas, we define

$$\chi(a) = 0, \qquad \text{if } (a, k) > 1.$$

This does not affect the validity of Theorem 6–6.

The duality of the relations of Theorem 6–6 is a reflection of the isomorphism of $X(k)$ and $M(k)$. In a sense, the reason for the importance of characters in the investigation of primes in progressions lies in the second relation, since it singles out the elements of a particular residue class (mod k), so that by use of the relation

$$\sum_{\substack{u \le a < v \\ a \equiv 1 \,(\text{mod } k)}} g(a) = \frac{1}{h} \sum_{u \le a < v} g(a) \sum_{\chi} \chi(a),$$

sums can be extended over an entire interval instead of a finite or infinite arithmetic progression $kt + 1$. Moreover, by a slight modification, any other residue class can be distinguished in the same way.

THEOREM 6–7. *If* $(a, k) = (b, k) = 1$, *then*

$$\sum_{\chi \in X(k)} \frac{\chi(a)}{\chi(b)} = \begin{cases} h & \text{if } a \equiv b \;(\text{mod } k), \\ 0 & \text{otherwise.} \end{cases}$$

Proof: Choose c so that $bc \equiv 1$ (mod k). Then

$$\sum_{\chi \in X(k)} \frac{\chi(a)}{\chi(b)} = \sum_{\chi \in X(k)} \chi(ac),$$

and, by Theorem 6–6, the last sum is h or 0 according as ac is or is not congruent to 1 (mod k), that is, according as a is or is not congruent to b (mod k).

It should be noticed that the function

$$\chi(n) = \begin{cases} (-1)^{\frac{1}{2}(n-1)} & \text{for } n \text{ odd,} \\ 0 & \text{for } n \text{ even,} \end{cases}$$

introduced in Section 6–1 is a character modulo 4. It and the principal character

$$\chi_0(n) = \begin{cases} 1 & \text{for } n \text{ odd,} \\ 0 & \text{for } n \text{ even,} \end{cases}$$

constitute the group $X(4)$ of order $\varphi(4) = 2$. The correspondence

$$\chi_0 \leftrightarrow 1, \qquad \chi \leftrightarrow 3,$$

describes the isomorphism between $X(4)$ and $M(4)$; each is the cyclic group of two elements.

6–3 The *L*-functions. For each character χ, we define a function $L(s, \chi)$ for $s > 1$ by the equation

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s},$$

or equivalently (according to Theorem 6–3) by the equation

$$L(s, \chi) = \prod_p \left(1 - \frac{\chi(p)}{p^s}\right)^{-1}. \tag{11}$$

In particular,

$$L(s, \chi_0) = \prod_{p \nmid k} (1 - p^{-s})^{-1} = \prod_{p \mid k} (1 - p^{-s})\zeta(s),$$

so that, by equation (2),

$$L(s, \chi_0) = \prod_{p \mid k} (1 - p^{-s}) \left(\frac{s}{s-1} - s\int_1^{\infty} \frac{(t)}{t^{s+1}}\,dt\right). \tag{12}$$

This latter representation for $L(s, \chi_0)$ is consistent with the series definition for $s > 1$, and may be taken as the definition for $0 < s < 1$.

For the proof of Dirichlet's theorem, it is necessary to know some of the properties of these *L*-functions. All the relevant properties can be proved by elementary arguments, but the proofs frequently can be simplified considerably if use is made of the theory of functions of a complex variable. In these cases alternative proofs will be given.

THEOREM 6–8. *$L(s, \chi_0)$ is continuous for $s > 1$, and*

$$\lim_{s \to 1^+} (s - 1)L(s, \chi_0) = \frac{h}{k}.$$

Proof: For $s \geq s_0 > 1$,

$$L(s, \chi_0) = \sum_{(n,k)=1} \frac{1}{n^s} \ll \sum_{n=1}^{\infty} \frac{1}{n^{s_0}} = \zeta(s_0),$$

so that the series for $L(s, \chi_0)$ converges uniformly in any interval to the right of $s = 1$. Since the separate terms are continuous, the sum is also continuous. Moreover, by (12),

$$\lim_{s \to 1^+} (s - 1)L(s, \chi_0) = \prod_{p \mid k} (1 - p^{-1}) = \frac{\varphi(k)}{k} = \frac{h}{k}.$$

For $\chi \neq \chi_0$, Theorem 6–6 shows that for arbitrary n_0,

$$\sum_{n=n_0}^{n_0+k} \chi(n) = 0,$$

so that by grouping the terms of

$$\sum_{n=u}^{v} \chi(n)$$

in blocks of k, with perhaps part of a block left over, we see that

$$\left| \sum_{n=u}^{v} \chi(n) \right| \leq h.$$

It follows from Theorem 6-2 that the Dirichlet series for $L(s, \chi)$ is convergent for $s > 0$. We need a slightly stronger result, which is proved in the following theorem.

THEOREM 6-9. *If $\chi \neq \chi_0$, then $L(s, \chi)$ has a continuous derivative (and is therefore itself continuous) for $s > 0$.*

Elementary proof: We use the standard theorem from analysis, that if the series resulting from termwise differentiation of a given series converges uniformly over an interval, then its sum is the derivative of the original series. The termwise derivative of

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

is

$$-\sum_{n=1}^{\infty} \frac{\chi(n) \log n}{n^s}, \tag{13}$$

and for $0 < s_0 \leq s \leq s_1$, the result follows from Theorem 6-2 by taking $a_n = \chi(n)$, $b_n(s) = n^{-s} \log n$. But s_0 may be arbitrarily small, and s_1 arbitrarily large, so that every $s > 0$ can be included in an interval in which $L(s, \chi)$ is continuously differentiable.

Alternative proof: Applying Theorem 6-2 and the fact that

$$\sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} \ll \sum_{n=1}^{\infty} \frac{1}{n^\sigma},$$

where $\sigma = \mathrm{Re}\, s$, we see that the series for $L(s, \chi)$ is uniformly convergent for $\mathrm{Re}\, s \geq \sigma_0 > 0$. Since each term of the series is an analytic function of s, the sum is also analytic, and is therefore differentiable.

6-4 Nonelementary proof of Dirichlet's theorem. There is a proof of Dirichlet's theorem which is remarkably simple and illuminating, and which fails to be elementary only in the sense that logarithms of

complex numbers are used. If the student who is not familiar with this extension will assume that the usual properties of logarithms of positive numbers (including the form of the Maclaurin expansion of $\log (1 + x)$ for $|x| < 1$) carry over to logarithms of nonzero complex numbers, he will find this proof much more straightforward than the elementary proof given in the following section, where use is made of the relation

$$\frac{d}{dx} \log f(x) = \frac{f'(x)}{f(x)}$$

to avoid logarithms entirely.

For $s > 1$, $|x(p)/p^s| < 1$, so that for such s we can describe a branch of the function $\log (1 - x(p)/p^s)$ by the equation

$$\log \left(1 - \frac{x(p)}{p^s}\right) = - \sum_{m=1}^{\infty} \frac{1}{m} \left(\frac{x(p)}{p^s}\right)^m = - \sum_{m=1}^{\infty} \frac{x(p^m)}{mp^{ms}}.$$

By (11), this induces the choice

$$\log L(s, x) = \sum_{p} \sum_{m=1}^{\infty} \frac{x(p^m)}{mp^{ms}}, \qquad \text{for } s > 1. \qquad (14)$$

THEOREM 6–10. *For each x, the function*

$$F(s, x) = \log L(s, x) - \sum_{p} \frac{x(p)}{p^s} \qquad (15)$$

is bounded in absolute value for $s \geqq 1$.

Proof: We rewrite (14) in the form

$$\log L(s, x) = \sum_{p} \frac{x(p)}{p^s} + \sum_{p} \sum_{m=2}^{\infty} \frac{x(p^m)}{mp^{ms}}.$$

Here,

$$\sum_{p} \sum_{m=2}^{\infty} \frac{x(p^m)}{mp^{ms}} \ll \sum_{p} \sum_{m=2}^{\infty} \frac{1}{2p^{ms}} = \frac{1}{2} \sum_{p} \frac{1}{p^{2s}(1 - p^{-s})}$$

$$\ll \frac{1}{2} \sum_{p} \frac{1}{p^{2s}(1 - 2^{-s})} = \frac{(1 - 2^{-s})^{-1}}{2} \sum_{p} \frac{1}{p^{2s}}$$

$$\ll \frac{(1 - 2^{-s})^{-1}}{2} \sum_{n=1}^{\infty} \frac{1}{n^{2s}} = \frac{(1 - 2^{-s})^{-1}}{2} \zeta(2s),$$

and since $\zeta(2s)$ is bounded for $2s \geqq 1 + \epsilon$, the theorem follows.

We can now complete the proof of Dirichlet's theorem, except for one gap which will be considered later.

THEOREM 6–11 (*Dirichlet's theorem*). *If $(k, l) = 1$, then there are infinitely many primes of the form $kt + l$.*

Proof: Multiply equation (15) by $1/\chi(l)$ and sum over all χ in $X(k)$. This gives

$$\sum_\chi \frac{\log L(s, \chi)}{\chi(l)} = \sum_\chi \sum_p \frac{\chi(p)}{\chi(l)p^s} + \sum_\chi \frac{F(s, \chi)}{\chi(l)}$$

$$= \sum_p \frac{1}{p^s} \sum_\chi \frac{\chi(p)}{\chi(l)} + \sum_\chi \frac{F(s, \chi)}{\chi(l)},$$

and, by Theorem 6–7,

$$\sum_\chi \frac{\log L(s, \chi)}{\chi(l)} = h \sum_{p \equiv l \,(\mathrm{mod}\, k)} \frac{1}{p^s} + \sum_\chi \frac{F(s, \chi)}{\chi(l)}. \tag{16}$$

Let $s \to 1^+$ in (16). The second term on the right remains bounded, by Theorem 6–10. We know that

$$\lim_{s \to 1^+} L(s, \chi_0) = \infty,$$

so that

$$\lim_{s \to 1^+} \frac{\log L(s, \chi_0)}{\chi_0(l)} = \infty.$$

Suppose for the moment that it had been shown that the remaining functions $L(s, \chi)$ (which we know to be continuous at $s = 1$) have nonzero values $L(1, \chi)$ at $s = 1$. It would follow that

$$\left| \lim_{s \to 1^+} \sum_{\chi \neq \chi_0} \frac{\log L(s, \chi)}{\chi(l)} \right| < \infty,$$

and (16) would then imply that

$$\lim_{s \to 1^+} \sum_{p \equiv l \,(\mathrm{mod}\, k)} \frac{1}{p^s} = \infty,$$

an equation which is possible only if the sum has infinitely many terms. Thus when we show that $L(1, \chi) \neq 0$ if $\chi \neq \chi_0$, we shall have proved not only Dirichlet's theorem but the stronger result that the series

$$\sum_{p \equiv l \,(\mathrm{mod}\, k)} \frac{1}{p}$$

diverges.

6–5 Elementary proof of Dirichlet's theorem. It is possible to avoid the complex logarithm $\log L(s, \chi)$ by using its derivative instead:

$$\frac{d}{dx} \log L(s, \chi) = \frac{L'(s, \chi)}{L(s, \chi)} = \frac{L'}{L}(s, \chi).$$

If we could use the relation (14), we could immediately deduce that

$$\frac{L'(s, \chi)}{L(s, \chi)} = - \sum_{p} \sum_{m=1}^{\infty} \frac{\chi(p^m) \log p}{p^{ms}};$$

since we cannot, we arrive at the same result by the rather more awkward method of dividing $L'(s, \chi)$ by $L(s, \chi)$. In the process, we shall have occasion to use some properties of the Möbius μ-function, which is defined by the following relations:

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n \text{ is divisible by a square larger than 1,} \\ (-1)^r & \text{if } n \text{ is the product of } r \text{ distinct primes.} \end{cases}$$

Alternatively, μ is, the multiplicative function (that is, $\mu(mn) = \mu(m)\mu(n)$ whenever $(m, n) = 1$) such that

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1, \\ -1 & \text{if } n = p, \text{ a prime,} \\ 0 & \text{if } n = p^\alpha, \quad \alpha > 1. \end{cases}$$

The properties we shall need are these.*

(a) $\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases}$

(b) If f is any number-theoretic function and

$$F(n) = \sum_{d|n} f(d),$$

then

$$f(n) = \sum_{d|n} \mu(d) F\left(\frac{n}{d}\right).$$

THEOREM 6–12. *If f is a completely multiplicative function, and the series*

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

---

* See, for example, Volume I, Theorems 6–5 and 6–6.

converges absolutely for $s > s_0$, then

$$\left(\sum_{n=1}^{\infty} \frac{f(n)}{n^s}\right)^{-1} = \sum_{n=1}^{\infty} \frac{f(n)\mu(n)}{n^s}$$

for $s > s_0$.

Proof: We have

$$\sum_{m=1}^{\infty} \frac{f(m)}{m^s} \sum_{n=1}^{\infty} \frac{f(n)\mu(n)}{n^s} = \sum_{m,n=1}^{\infty} \frac{f(mn)\mu(n)}{(mn)^s}$$

$$= \sum_{j=1}^{\infty} \frac{\sum_{d|j} \mu(d)}{j^s} f(j) = 1.$$

THEOREM 6–13.   *For each χ, the relation*

$$\frac{L'}{L}(s, \chi) = -\sum_{n=1}^{\infty} \frac{\chi(n)\Lambda(n)}{n^s} \tag{17}$$

holds for $s > 1$, where

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^\alpha \text{ for some } \alpha > 0 \text{ and prime } p, \\ 0 & \text{otherwise.} \end{cases}$$

Proof: By the preceding theorem and the expression (13) for $L'(s, \chi)$, we have, for $s > 1$,

$$\frac{L'}{L}(s, \chi) = -\sum_{n=1}^{\infty} \frac{\chi(m)\log m}{m^s} \cdot \sum_{j=1}^{\infty} \frac{\chi(j)\mu(j)}{j^s}$$

$$= -\sum_{m,j=1}^{\infty} \frac{\chi(mj)\mu(j)\log m}{(mj)^s}$$

$$= -\sum_{n=1}^{\infty} \frac{\chi(n) \sum_{d|n} \mu(d) \log \frac{n}{d}}{n^s}.$$

But from the ob... 's relation

$$\log n = \sum_{d|n} \Lambda(d)$$

and the Möbius inversion formula quoted above, we have

$$\Lambda(n) = \sum_{d|n} \mu(d) \log \frac{n}{d},$$

and the theorem follows.

THEOREM 6–14. *For each χ, the function*

$$G(s, \chi) = \frac{L'}{L}(s, \chi) + \sum_p \chi(p) \frac{\log p}{p^s} \tag{18}$$

is bounded in absolute value for $s \geq 1$.

Proof: Equation (17) may be rewritten in the form

$$\frac{L'}{L}(s, \chi) = -\sum_p \frac{\chi(p) \log p}{p^s} - \sum_p \sum_{m=2}^{\infty} \frac{\chi(p^m) \log p}{p^{ms}},$$

and

$$\sum_p \sum_{m=2}^{\infty} \frac{\chi(p^m) \log p}{p^{ms}} \ll \sum_p \sum_{m=2}^{\infty} \frac{\log p}{p^{ms}} = \sum_p \frac{\log p}{p^{2s}(1 - p^{-s})}$$

$$\ll \sum_p \frac{\log p}{p^{2s}(1 - 2^{-s})},$$

and the last series clearly converges for $s > \frac{1}{2}$.

We can now complete the proof of Dirichlet's theorem in much the same way as before. Multiplying both sides of (18) by $1/\chi(l)$, and summing over all χ in $X(k)$, we obtain

$$\sum \frac{1}{\chi(l)} \frac{L'}{L}(s, \chi) = -\sum_\chi \sum_p \frac{\chi(p)}{\chi(l)} \frac{\log p}{p^s} + \sum_\chi \frac{1}{\chi(l)} G(s, \chi)$$

$$= -h \sum_{p \equiv l \,(\mathrm{mod}\, k)} \frac{\log p}{p^s} + \sum_\chi \frac{1}{\chi(l)} G(s, \chi).$$

Now let $s \to 1^+$. The second term on the right remains bounded. Assuming again that $L(1, \chi) \neq 0$ for $\chi \neq \chi_0$, the quantity $1/L(s, \chi)$ is also bounded for s sufficiently close to 1, since L is continuous at $s = 1$. For $\chi \neq \chi_0$, $L'(s, \chi)$ remains bounded, by Theorem 6–9. On the other hand,

$$\left| \frac{L'}{L}(s, \chi_0) \right| = \sum_{(n,k)=1} \frac{\Lambda(n)}{n^s} = \sum_{p \nmid k} \frac{\log p}{p^s} > \sum_{p \nmid k} \frac{1}{p^s}$$

$$= \log L(s, \chi_0) + F(s, \chi_0),$$

by Theorem 6–10, and the quantity $\log L(s, \chi_0) + F(s, \chi_0)$ increases without bound as $s \to 1^+$. It follows that

$$\lim_{s \to 1^+} \sum_{p \equiv l \,(\mathrm{mod}\, k)} \frac{\log p}{p^s} = \infty,$$

and the theorem is again proved, except for the verification of the fact that $L(1, \chi) \neq 0$ for $\chi \neq \chi_0$.

6–6 Proof that $L(1, \chi) \neq 0$

THEOREM 6–15. *If χ assumes a nonreal value for some n, then $L(1, \chi) \neq 0$.*

Proof: Let χ be such a character, and let $\bar{\chi}$ be the function whose value for each a is the complex conjugate of that of χ. Clearly $\bar{\chi}$ is also a character, and $\bar{\chi} \neq \chi$. But if $L(1, \chi) = 0$, then also

$$L(1, \bar{\chi}) = \overline{L(1, \chi)} = 0,$$

so at least two L-functions must vanish in this case. Since $L(s, \chi)$ is differentiable at $s = 1$, the quantities

$$L'(1, \chi) = \lim_{s \to 1} \frac{L(s, \chi)}{s - 1} \quad \text{and} \quad L'(1, \bar{\chi}) = \lim_{s \to 1} \frac{L(s, \bar{\chi})}{s - 1}$$

exist, so that there is a number A such that

$$\lim_{s \to 1^+} \frac{\prod\limits_{\chi \neq \chi_0} L(s, \chi)}{(s - 1)^2} = A.$$

Since

$$\lim_{s \to 1^+} (s - 1)L(s, \chi_0) = \frac{h}{k},$$

we deduce that

$$\lim_{s \to 1^+} \prod_\chi L(s, \chi) = \lim_{s \to 1^+} \left\{ (s - 1)\big((s - 1)L(s, \chi_0)\big) \frac{\prod\limits_{\chi \neq \chi_0} L(s, \chi)}{(s - 1)^2} \right\}$$

$$= 0 \cdot \frac{h}{k} \cdot A = 0.$$

But by (14),

$$\sum_\chi \log L(s, \chi) = \sum_\chi \sum_p \sum_{m=1}^\infty \frac{\chi(p^m)}{m p^{ms}}$$

$$= \sum_p \sum_{m=1}^\infty \frac{\sum\limits_\chi \chi(p^m)}{m p^{ms}}$$

$$= h \sum_{\substack{p,m \\ p^m \equiv 1 \,(\text{mod } k)}} \frac{1}{m p^{ms}} > 0$$

for $s > 1$, so that

$$\lim_{s \to 1^+} \prod_{\chi} L(s, \chi) > e^0 = 1.$$

This contradiction establishes the theorem.

It would not be easy to avoid the use of the complex logarithm in this proof, since the Dirichlet series for $\prod_{\chi} L(s, \chi)$ has very complicated coefficients. To obtain an elementary proof, it is simpler to use a different combination of L-functions. Unfortunately, the choice we make can hardly be motivated by an elementary argument, but must remain a *deus ex machina* until Section 7-3. It is the left side of the inequality

$$L^3(s, \chi_0)|L(s, \chi)|^4|L(s, \chi^2)|^2 \geq 1; \tag{19}$$

this inequality we now show to be valid for $s > 1$.

Note first that for $z = r(\cos\theta + i\sin\theta)$,

$$|1 - z|^2 = |1 - r\cos\theta - ir\sin\theta|^2 = 1 - 2r\cos\theta + r^2,$$

and that for arbitrary real θ,

$$2\cos\theta + \cos 2\theta = 2\cos\theta + 2\cos^2\theta - 1 = 2(\cos\theta + \tfrac{1}{2})^2 - \tfrac{3}{2} \geq -\tfrac{3}{2}.$$

Using the fact that the geometric mean of three positive numbers is at most equal to their arithmetic mean, we see that, if $p \nmid k$ and

$$\chi(p) = \cos\theta_p + i\sin\theta_p,$$

then

$$\left(\left| 1 - \frac{\chi(p)}{p^s} \right|^2 \right)^2 \left| 1 - \frac{\chi^2(p)}{p^s} \right|^2$$

$$= (1 - 2p^{-s}\cos\theta_p + p^{-2s})^2(1 - 2p^{-s}\cos 2\theta_p + p^{-2s})$$

$$\leq (1 - \tfrac{2}{3}p^{-s}(2\cos\theta_p + \cos 2\theta_p) + p^{-2s})^3$$

$$\leq (1 + p^{-s} + p^{-2s})^3 \leq \left(\frac{1}{1 - p^{-s}} \right)^3,$$

or

$$(1 - \chi_0(p)p^{-s})^3|1 - \chi(p)p^{-s}|^4|1 - \chi^2(p)p^{-s}|^2 \leq 1.$$

This inequality also holds if $p \mid k$, and, multiplying over all p, we obtain (19).

It is now simple to prove that $L(1, \chi) \neq 0$ if $\chi^2 \neq \chi_0$, that is, if χ is nonreal. Supposing the opposite, and using the fact that $L'(s, \chi)$ is

continuous at $s = 1$, we have that for $1 \leq s \leq s_1$,

$$|L(s, \chi)| = |L(s, \chi) - L(1, \chi)| = \left| \int_1^s L'(u, \chi)\, du \right| \leq A_1(s - 1),$$

where $A_1 = \max_{1 \leq s \leq s_1} |L'(s, \chi)|.$

But now (19) can be recast in the form

$$(s - 1)((s - 1)L(s, \chi_0))^3 \left| \frac{L(s, \chi)}{s - 1} \right|^4 |L(s, \chi^2)|^2 \geq 1,$$

in which the first factor tends to zero and the others remain bounded, as $s \to 1^+$. This inequality is false for some $s > 1$, and the contradiction shows that $L(1, \chi) \neq 0$.

No device of this sort has been found for the case that $\chi(n)$ is real for all n. Showing that $L(1, \chi) \neq 0$ for a real character is the most difficult point in the entire proof. Dirichlet effected it by showing that $L(1, \chi)$ is a factor in the class number of a certain quadratic field. This and other algebraic proofs require a considerable amount of background; we shall content ourselves with an elementary and a function-theoretic proof. We first sketch the idea.

If $s > 1$, then

$$\zeta(s)L(s, \chi) = \sum_{m=1}^{\infty} \frac{1}{m^s} \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s} = \sum_{m,n=1}^{\infty} \frac{\chi(n)}{(mn)^s} = \sum_{t=1}^{\infty} \frac{\sum_{n|t} \chi(n)}{t^s},$$

so that if we put

$$f(n) = \sum_{d|n} \chi(d), \tag{20}$$

then $$\zeta(s)L(s, \chi) = \sum_{n=1}^{\infty} \frac{f(n)}{n^s} \tag{21}$$

for $s > 1$.

By Theorem 6–17, below,

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} \geq \sum_{m=1}^{\infty} \frac{1}{(m^2)^s} = \zeta(2s),$$

so that even in   series $\sum f(n)n^{-s}$ converges to the right of $s = \frac{1}{2}$, it is certainly not bounded near $s = \frac{1}{2}$. In the analytic proof, we show that (21) is correct for $s \geq \frac{1}{2}$ if $L(1, \chi) = 0$, and obtain the contradiction

$$\lim_{s \to \frac{1}{2}^+} L(s, \chi)\zeta(s) = L(\tfrac{1}{2}, \chi)\zeta(\tfrac{1}{2}) = \infty.$$

In the elementary proof, questions of convergence are avoided by considering partial sums for $s = \frac{1}{2}$ rather than the full series for s near $\frac{1}{2}$. It will be shown that

$$\sum_{n=1}^{x} \frac{f(n)}{\sqrt{n}} = 2\sqrt{x}\, L(1, \chi) + O(1),$$

and also that the sum on the left tends to infinity with x, so that the relation $L(1, \chi) = 0$ is impossible.

THEOREM 6–16. *With f as in* (20),

$$f(n) \geq \begin{cases} 0 & \textit{for all } n, \\ 1 & \textit{for square } n. \end{cases}$$

Proof: Being the arithmetic sum function of a multiplicative function, f is itself multiplicative,* so that

$$f(p_1^{\alpha_1} \cdots p_r^{\alpha_r}) = f(p_1^{\alpha_1}) \cdots f(p_r^{\alpha_r}).$$

Since χ is a real character, $\chi(p) = 0$ or ± 1 for each prime p, and

$$f(p^\alpha) = \sum_{\beta=0}^{\alpha} \chi(p^\beta)$$

$$= \sum_{\beta=0}^{\alpha} (\chi(p))^\beta = \begin{cases} 1 + 0 + \cdots + 0 & \text{if } \chi(p) = 0, \\ 1 + 1 + \cdots + 1 & \text{if } \chi(p) = 1, \\ 1 - 1 + \cdots + (-1)^\alpha & \text{if } \chi(p) = -1. \end{cases}$$

Hence

$$f(p^\alpha) = \begin{cases} 1 & \text{if } \chi(p) = 0, \\ \alpha + 1 & \text{if } \chi(p) = 1, \\ 1 & \text{if } \chi(p) = -1, \alpha \text{ even}, \\ 0 & \text{if } \chi(p) = -1, \alpha \text{ odd}, \end{cases}$$

and the theorem follows.

THEOREM 6–17. *The relations*

$$\sum_{n=1}^{x} \chi(n) = O(1) \tag{22}$$

and

$$\sum_{n=x}^{\infty} \frac{\chi(n)}{n^s} = O\left(\frac{1}{x^s}\right) \tag{23}$$

hold as $x \to \infty$, *for* $s > 0$.

* Cf. Volume I, Theorem 6–3.

Proof: We have already noticed that if

$$S(x) = \sum_{n=1}^{x} \chi(n),$$

then $|S(x)| \leq h$, which implies (22). Using this, we have

$$\left| \sum_{n=x}^{\infty} \frac{\chi(n)}{n^s} \right| = \left| \sum_{n=x}^{\infty} \frac{S(n) - S(n-1)}{n^s} \right|$$

$$= \left| \sum_{n=x}^{\infty} S(n) \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) - \frac{S(x-1)}{x^s} \right|$$

$$\leq h \sum_{n=x}^{\infty} \left(\frac{1}{n^s} - \frac{1}{(n+1)^s} \right) + \frac{h}{x^s} = \frac{2h}{x^s}$$

which implies (23).

THEOREM 6-18.   *There is a constant C such that*

$$\sum_{n=1}^{x} \frac{1}{\sqrt{n}} = 2\sqrt{x} + C + O\left(\frac{1}{\sqrt{x}} \right).$$

Proof: Put

$$t_n = 2\sqrt{n} - 2\sqrt{n-1} - \frac{1}{\sqrt{n}} = \int_{n-1}^{n} \frac{dx}{\sqrt{x}} - \frac{1}{\sqrt{n}},$$

so that

$$\sum_{n=2}^{x} t_n = 2\sqrt{x} - 2 - \sum_{n=2}^{x} \frac{1}{\sqrt{n}}.$$

Now t_n, being the area of the triangular region bounded by the curve $y = x^{-\frac{1}{2}}$ and the lines $x = n - 1$ and $y = n^{-\frac{1}{2}}$, is positive and smaller than $(n-1)^{-\frac{1}{2}} - n^{-\frac{1}{2}}$, so that the series

$$\sum_{n=1}^{\infty} t_n$$

converges, and

$$\sum_{n=x+1}^{\infty} t_n < \sum_{n=x+1}^{\infty} \left(\frac{1}{\sqrt{n-1}} - \frac{1}{\sqrt{n}} \right) = \frac{1}{\sqrt{x}}.$$

Hence

$$2\sqrt{x} - \sum_{n=1}^{x} \frac{1}{\sqrt{n}} = 1 + \sum_{n=2}^{\infty} t_n - \sum_{n=x+1}^{\infty} t_n = 1 + \sum_{n=2}^{\infty} t_n + O\left(\frac{1}{\sqrt{x}} \right).$$

This proves the theorem for integral x; its extension to real x is immediate.

THEOREM 6–19. *If $\chi \neq \chi_0$ is a real character, then $L(1, \chi) \neq 0$.*

Proof: Put

$$G(x) = \sum_{n=1}^{x} \frac{f(n)}{\sqrt{n}} .$$

By Theorem 6–16,

$$G(x) \geq \sum_{m=1}^{\sqrt{x}} \frac{1}{\sqrt{m^2}} = \sum_{m=1}^{\sqrt{x}} \frac{1}{m} ,$$

so that $G(x) \to \infty$ with x.

On the other hand,

FIGURE 6–1

$$G(x) = \sum_{j=1}^{x} \frac{1}{\sqrt{j}} \sum_{d|j} \chi(d) = \sum_{uv \leq x} \frac{\chi(v)}{\sqrt{uv}} .$$

This sum, extended over the lattice points u, v for which $u \geq 1$, $v \geq 1$, $uv \leq x$, we split into two parts, as indicated in Fig. 6–1:

$$G(x) = \sum_{u=1}^{\sqrt{x}} \sum_{v=\sqrt{x}+1}^{x/u} \frac{\chi(v)}{\sqrt{uv}} + \sum_{v=1}^{\sqrt{x}} \sum_{u=1}^{x/v} \frac{\chi(v)}{\sqrt{uv}}$$

$$= \sum_{u=1}^{\sqrt{x}} \frac{1}{\sqrt{u}} \sum_{v=\sqrt{x}+1}^{x/u} \frac{\chi(v)}{\sqrt{v}} + \sum_{v=1}^{\sqrt{x}} \frac{\chi(v)}{\sqrt{v}} \sum_{u=1}^{x/v} \frac{1}{\sqrt{u}}$$

$$= \sum_{u=1}^{\sqrt{x}} \frac{1}{\sqrt{u}} O(x^{-\frac{1}{4}}) + \sum_{v=1}^{\sqrt{x}} \frac{\chi(v)}{\sqrt{v}} \left(2\sqrt{\frac{x}{v}} + C + O\left(\sqrt{\frac{v}{x}}\right) \right)$$

$$= O(x^{-\frac{1}{4}}) \cdot O(x^{\frac{1}{4}}) + 2\sqrt{x} \sum_{v=1}^{\sqrt{x}} \frac{\chi(v)}{v} + C \cdot O(1) + \frac{1}{\sqrt{x}} O(1)$$

$$= 2\sqrt{x} \left(L(1, \chi) + O\left(\frac{1}{\sqrt{x}}\right) \right) + O(1),$$

so that

$$\sum_{n=1}^{x} \frac{f(n)}{\sqrt{n}} = 2\sqrt{x} \, L(1, \chi) + O(1). \tag{24}$$

Thus, if $L(1, \chi)$ were zero, $G(x)$ would remain bounded as $x \to \infty$, which is not the case.

A rather more straightforward (function-theoretic) proof can be obtained by extending (21), which we know to be valid for $s > 1$, to
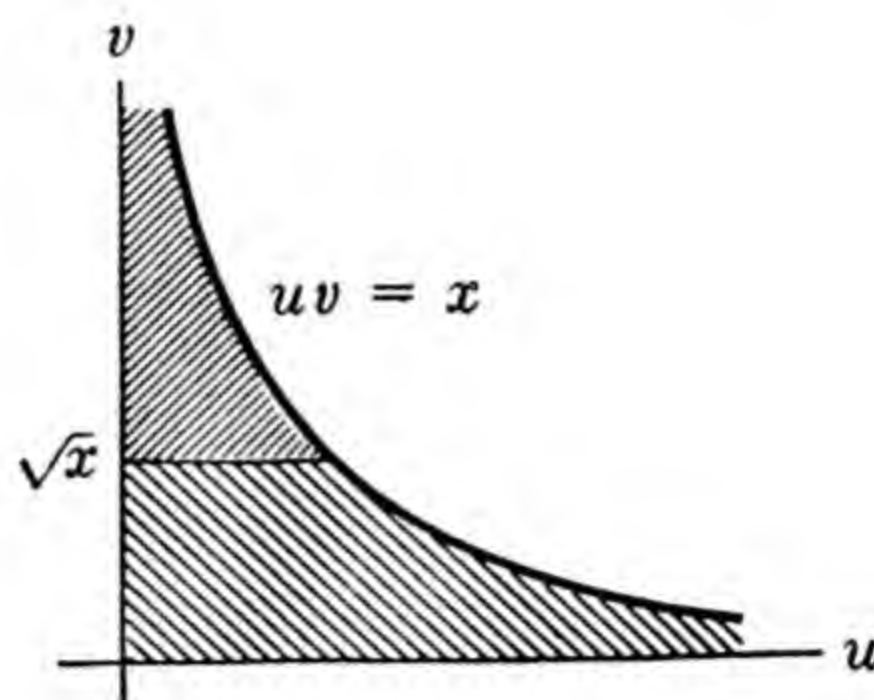
the range $s > \frac{1}{2}$, under the assumption that $L(1, \chi) = 0$. By an argument quite similar to, but slightly simpler than that which yielded (24), it can be shown that if $L(1, \chi) = 0$, then

$$\sum_{n=1}^{x} f(n) = O(\sqrt{x}).$$

Theorem 6–2 implies that the series

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s} = \sum_{n=1}^{\infty} \frac{f(n)/n^{\frac{1}{2}}}{n^{s-\frac{1}{2}}} = \sum_{n=1}^{\infty} \frac{O(1)}{n^{s-\frac{1}{2}}}$$

converges for $s > \frac{1}{2}$. Now let σ_0 be a real number greater than $\frac{1}{2}$, and let s be a complex number with $\operatorname{Re} s = \sigma > \sigma_0$. Then for $v > u > 1$, we have

$$\sum_{n=u}^{v} \frac{f(n)}{n^s} = \sum_{n=u}^{v} \frac{f(n)/n^{\sigma_0}}{n^{s-\sigma_0}}$$

$$= \sum_{n=u}^{v} \frac{\sum_{m=1}^{n} \frac{f(m)}{m^{\sigma_0}} - \sum_{m=1}^{n-1} \frac{f(m)}{m^{\sigma_0}}}{n^{s-\sigma_0}}$$

$$= \sum_{n=u}^{v-1} \sum_{m=1}^{n} \frac{f(m)}{m^{\sigma_0}} \left(\frac{1}{n^{s-\sigma_0}} - \frac{1}{(n+1)^{s-\sigma_0}} \right)$$

$$+ v^{-(s-\sigma_0)} \sum_{m=1}^{v} \frac{f(m)}{m^{\sigma_0}} - u^{-(s-\sigma_0)} \sum_{m=1}^{u-1} \frac{f(m)}{m^{\sigma_0}},$$

so that

$$\left| \sum_{n=u}^{v} \frac{f(n)}{n^s} \right|$$

$$\leq A \sum_{n=u}^{v-1} \left| \frac{1}{n^{s-\sigma_0}} - \frac{1}{(n+1)^{s-\sigma_0}} \right| + A|v^{-(s-\sigma_0)}| + A|u^{-(s-\sigma_0)}|$$

$$= A \sum_{n=u}^{v-1} \left| (s - \sigma_0) \int_{n}^{n+1} \frac{dz}{z^{s-\sigma_0+1}} \right| + Av^{-(\sigma-\sigma_0)} + Au^{-(\sigma-\sigma_0)}$$

$$\leq A \sum_{n=u}^{v-1} \frac{1}{n^{\sigma-\sigma_0+1}} + Av^{-(\sigma-\sigma_0)} + Au^{-(\sigma-\sigma_0)},$$

where A is such that

$$\left| \sum_{n=1}^{N} \frac{f(n)}{n^{\sigma_0}} \right| < A, \qquad \text{for all } N \geq 1.$$

It follows that the series

$$\sum_{n=1}^{\infty} \frac{f(n)}{n^s}$$

converges to an analytic function in the half-plane $\sigma > \frac{1}{2}$. Since it coincides with $L(s, \chi)\zeta(s)$ for s real and larger than 1, it represents an analytic continuation of $L(s, \chi)\zeta(s)$ for $\sigma > \frac{1}{2}$. But it is unbounded near $s = \frac{1}{2}$, while $L(s, \chi)\zeta(s)$ is not.

CHAPTER 7

THE PRIME NUMBER THEOREM

7-1 Introduction. It is shown in elementary number theory texts* that if

$$\lim_{x \to \infty} \frac{\pi(x)}{x/\log x}$$

exists, it must have the value 1, and that there are positive constants c and c' such that for $x \geq 2$,

$$c < \frac{\pi(x)}{x/\log x} < c'.$$

Neither of these results implies the other, of course; together they show that

$$0 < \liminf_{x \to \infty} \frac{\pi(x)}{x/\log x} \leq 1 \leq \limsup_{x \to \infty} \frac{\pi(x)}{x/\log x} < \infty.$$

(Here, as always, $\pi(x)$ denotes the number of primes less than or equal to x.) Both results were obtained by Chebyshev in 1851 and 1852 (in rather more precise form), but it was not until some forty-five years later that the final link was supplied by Hadamard and de la Vallée Poussin, who showed independently that the limit actually exists, and thus proved the Prime Number Theorem. Both proofs made essential use of the theory of functions of a complex variable, and despite much effort it seemed for many years impossible to give a proof entirely free of considerations as sophisticated as this theory. In 1948, however, P. Erdös and A. Selberg gave a completely elementary proof. More precisely, Selberg proved the fundamental relation

$$\sum_{p \leq x} \log^2 p + \sum_{pq \leq x} \log p \log q = 2x \log x + O(x),$$

and he and Erdös deduced the Prime Number Theorem from it.†

---

* See, for example, Volume I, Sections 6-6 and 6-7.

† Excellent expositions of this proof are given in T. Nagell, *Introduction to Number Theory* (New York: John Wiley & Sons, 1951) and in G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (3rd edition, New York: Oxford University Press, 1954).

We present a proof based on the behavior of the ζ-function for complex s. Throughout this chapter, familiarity with the contents of a standard course in the theory of functions of a complex variable is presupposed.

Before going into detail, we outline the proof. Our object is to get an estimate for

$$\pi(x) = \sum_{p \le x} 1 = \sum_{n=1}^{x} P(n),$$

where P is the characteristic function of the primes:

$$P(n) = \begin{cases} 1 & \text{if } n \text{ is prime,} \\ 0 & \text{otherwise.} \end{cases}$$

While P itself does not arise in a natural way, the function P^* such that

$$P^*(n) = \begin{cases} \dfrac{1}{m} & \text{if } n = p^m \text{ for some } m, p, \\ 0 & \text{otherwise,} \end{cases}$$

occurs in the Dirichlet series for $\log \zeta(s)$:

$$\log \zeta(s) = \sum_{m,p} \frac{1}{mp^{ms}} = \sum_{n=1}^{\infty} \frac{P^*(n)}{n^s}. \tag{1}$$

For fixed m, the number of mth powers of primes which do not exceed x is equal to the number of primes which do not exceed $\sqrt[m]{x}$, so that

$$\sum_{n=1}^{x} P^*(n) = \sum_{n=1}^{x} P(n) + \frac{1}{2} \sum_{n=1}^{\sqrt{x}} P(n) + \frac{1}{3} \sum_{n=1}^{\sqrt[3]{x}} P(n) + \cdots$$

$$= \pi(x) + \frac{\pi(\sqrt{x})}{2} + \frac{\pi(\sqrt[3]{x})}{3} + \cdots,$$

and since, for $m \ge 2$,

$$\pi(x^{1/m}) < cx^{1/m} \le c\sqrt{x} = o\left(\frac{x}{\log x}\right),$$

it is to be expected that

$$\sum_{n=1}^{x} P^*(n) \sim \pi(x).$$

In light of (1), the present case is a specialization of the following problem: given a function

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s},$$

to estimate

$$\sum_{n=1}^{x} a_n.$$ (2)

It will be shown that

$$J(w) = \frac{1}{2\pi i} \int_{2-\infty i}^{2+\infty i} \frac{c^{ws}}{s^2} \, ds = \begin{cases} w & \text{if } w \geq 0, \\ 0 & \text{if } w \leq 0, \end{cases}$$

so that $J(w)$ is closely related to the characteristic function of the positive real numbers. If we put

$$w = \log \frac{x}{n},$$

this gives

$$\frac{1}{2\pi i} \int_{2-\infty i}^{2+\infty i} \frac{(x/n)^s}{s^2} \, ds = \begin{cases} \log (x/n) & \text{if } n \leq x, \\ 0 & \text{if } n \geq x, \end{cases}$$

so that

$$\frac{1}{2\pi i} \int_{2-\infty i}^{2+\infty i} \frac{x^s}{s^2} f(s) \, ds = \sum_{n \leq x} a_n \log \frac{x}{n}.$$

If $\delta = \delta(x)$ tends monotonically to zero as $x \to \infty$, then

$$\sum_{n \leq x(1+\delta)} a_n \log \frac{x(1+\delta)}{n} - \sum_{n \leq x} a_n \log \frac{x}{n}$$

$$= \log (1 + \delta) \sum_{n \leq x} a_n + \sum_{x < n \leq x+\delta x} a_n \log \frac{x(1+\delta)}{n}$$

$$= \log (1 + \delta) \sum_{n \leq x} a_n + O \left(\log (1+\delta) \sum_{x < n \leq x+\delta x} a_n \right).$$

If the remainder term here is of smaller order of magnitude than the first term for suitable choice of δ, then

$$\sum_{n \leq x} a_n \sim \frac{1}{2\pi i \delta} \int_{2-\infty i}^{2+\infty i} \frac{x^s((1+\delta)^s - 1)}{s^2} f(s) \, ds,$$

and the problem reduces to that of obtaining an adequate estimate of this integral. To do this, we replace the line of integration by a suitable large closed contour, inside and on which we have sufficient information about $f(s)$ to apply standard contour-integral techniques.

In the case at hand, the estimation of the integral in the last relation requires some knowledge of the zeros, poles, and size of $\zeta(s)$.

7–2 Preliminary results. Following the odd but harmless tradition in analytic number theory, we designate by σ and t the real and imaginary parts of the complex variable s. For $x > 0$, x^s means $e^{s \log x}$, where $\log x$ indicates the real logarithm.

When we have proved the Prime Number Theorem, we shall consider some other rather similar problems, and for one of these it will be necessary to use not the Riemann ζ-function but the so-called *Hurwitz ζ-function*, defined for $0 < w \le 1$ and $\sigma > 1$ by the equation

$$\zeta(s, w) = \sum_{n=0}^{\infty} \frac{1}{(n + w)^s}.$$

Since $\zeta(s, 1) = \zeta(s)$, and since the requisite properties are no more difficult to prove for $\zeta(s, w)$ than for $\zeta(s)$, we consider the more general function.

THEOREM 7–1. *For any $\sigma_0 > 1$, the series*

$$\sum_{n=0}^{\infty} (n + w)^{-s}$$

converges uniformly for $\sigma \ge \sigma_0$, so that $\zeta(s, w)$ is regular (or analytic) for $\sigma > 1$.

Proof: We have

$$|(n + w)^{-s}| = |e^{-(\sigma + it) \log (n+w)}| = e^{-\sigma \log (n+w)} = (n + w)^{-\sigma},$$

so that for $\sigma \ge \sigma_0$,

$$\sum_{n=0}^{\infty} (n + w)^{-s} \ll \sum_{n=0}^{\infty} (n + w)^{-\sigma_0}.$$

Thus we have a series of analytic functions which is dominated throughout the region $\sigma \ge \sigma_0$ by a convergent series of positive constants, and which is therefore uniformly convergent, and the result follows from Weierstrass' theorem.

THEOREM 7–2. *If a and b are integers with $b > a \ge 0$, and if f has a continuous derivative over $a \le x \le b$, then*

$$\sum_{n=a+1}^{b} f(n) = \int_a^b f(u)\, du + \int_a^b (u - [u])f'(u)\, du.$$

Proof: We have

$$\int_{n-1}^{n} uf'(u)\, du = nf(n) - (n-1)f(n-1) - \int_{n-1}^{n} f(u)\, du$$

$$= f(n) + (n-1)\int_{n-1}^{n} f'(u)\, du - \int_{n-1}^{n} f(u)\, du$$

$$= f(n) + \int_{n-1}^{n} [u]f'(u)\, du - \int_{n-1}^{n} f(u)\, du,$$

from which the result follows by summing on n from $a+1$ to b.

THEOREM 7-3. *If m is a non-negative integer, and $\sigma > 1$, then*

$$\zeta(s,w) - \frac{1}{(s-1)(m+w)^{s-1}} = \sum_{n=0}^{m} \frac{1}{(n+w)^s} - s\int_{m}^{\infty} \frac{u-[u]}{(u+w)^{s+1}}\, du. \quad (3)$$

It follows that $\zeta(s,w) - 1/(s-1)$ is regular for $\sigma > 0$, and that (3) holds for $\sigma > 0$.

Proof: If $\sigma > 1$ and

$$f(u) = \frac{1}{(u+w)^s},$$

then the equation of Theorem 7-2 continues to hold if $b \to \infty$, and, replacing a by m, we have

$$\sum_{n=m+1}^{\infty} \frac{1}{(n+w)^s} = \frac{1}{(s-1)(m+w)^{s-1}} - s\int_{m}^{\infty} \frac{u-[u]}{(u+w)^{s+1}}\, du,$$

from which (3) follows. Since

$$\left| \frac{u-[u]}{(u+w)^{s+1}} \right| < \frac{1}{(u+w)^{\sigma+1}} < \frac{1}{u^{\sigma+1}},$$

the integral on the right side of (3) converges absolutely for $\sigma > 0$, and uniformly for $\sigma \geq \sigma_0 > 0$. For arbitrary $n \geq 0$, the quantity

$$\int_{n}^{n+1} \frac{u-[u]}{(u+w)^{s+1}}\, du = \int_{n}^{n+1} \frac{u-n}{(u+w)^{s+1}}\, du$$

is a regular function of s for $\sigma > 0$; the same is true of

$$\sum_{n=m}^{\infty} \int_{n}^{n+1} \frac{u-[u]}{(u+w)^{s+1}}\, du = \int_{m}^{\infty} \frac{u-[u]}{(u+w)^{s+1}}\, du,$$

for $m \geq 0$. Finally, taking $m = 0$ in (3), we have

$$\zeta(s, w) - \frac{1}{s - 1} = \frac{1}{w^s} + \frac{w^{1-s} - 1}{s - 1} - s \int_0^\infty \frac{u - [u]}{(u + w)^{s+1}} \, du,$$

and the right side is regular for $\sigma > 0$.

Equation (3) thus provides an analytic continuation of $\zeta(s, w)$ over the half-plane $\sigma > 0$. The function is actually analytic over the entire plane, except for the pole at $s = 1$, but this fact is not needed.

Hereafter c will denote a positive constant which depends only on the arguments indicated; it need not have the same value in different occurrences, unless it has a subscript.

THEOREM 7–4. *For* $\frac{1}{2} \leq \sigma \leq 2$ *and* $t > c(w)$,

$$|\zeta(s, w)| < t^{\frac{3}{4}}.$$

For $t \geq 8$ *and* $1 - (\log t)^{-1} \leq \sigma \leq 2$,

$$|\zeta(s, w)| < c(w) \log t.$$

Proof: For $\frac{1}{2} \leq \sigma \leq 2$ and $t \geq 3$ we have $|s| < 2 + t < 2t$ and $|s - 1| \geq t > 1$. Hence if we take $m = [t] + 1$ in (3), we have

$$|\zeta(s, w)| < \frac{1}{([t] + 1 + w)^{\sigma - 1}} + \frac{1}{w^\sigma} + \sum_{n=1}^{[t]+1} \frac{1}{n^\sigma} + 2t \int_t^\infty \frac{du}{u^{\sigma+1}}$$

$$\leq \frac{1}{([t] + 1 + w)^{\sigma - 1}} + c(w) + \sum_{n=1}^{[t]} \frac{1}{n^\sigma} + \frac{2t}{\sigma t^\sigma},$$

or

$$|\zeta(s, w)| < \frac{1}{([t] + 1 + w)^{\sigma - 1}} + c(w) + \sum_{n=1}^{[t]} \frac{1}{n^\sigma} + 4t^{1-\sigma}. \qquad (4)$$

Thus, for this same range of σ and t,

$$|\zeta(s, w)| < \frac{1}{([t] + 1 + w)^{-\frac{1}{2}}} + c(w) + \sum_{n=1}^{[t]} \frac{1}{\sqrt{n}} + 4\sqrt{t}$$

$$< 2\sqrt{t} + c(w) + \int_0^t \frac{du}{\sqrt{u}} + 4\sqrt{t} \leq 8\sqrt{t} + c(w),$$

and this is smaller than $t^{\frac{3}{4}}$ for $t > c(w)$.

Now take $t \geq 8 > e^2$. Then $1 - (\log t)^{-1} \geq \frac{1}{2}$, so that if $1 - (\log t)^{-1} \leq \sigma \leq 2$, the inequality (4) gives

$$|\zeta(s, w)| < (2t)^{1/\log t} + c(w) + \sum_{n=1}^{[t]} \frac{n^{1/\log t}}{n} + 4i^{1/\log t}$$

$$< 2^{\frac{1}{2}}e + c(w) + e \sum_{n=1}^{[t]} \frac{1}{n} + 4e < c(w) \log t.$$

THEOREM 7–5. *If, for $|x| \leq 1$,*

$$f(x) = \sum_{n=1}^{\infty} a_n x^n$$

is regular and Re $f(x) \leq \frac{1}{2}$, *then* $|a_n| \leq 1$ *for* $n \geq 1$.

Proof: Since $|f(x)| \leq |1 - f(x)|$ for $|x| \leq 1$, the function

$$\frac{f(x)}{1 - f(x)} = \frac{a_1 x + \cdots}{1 - a_1 x - \cdots} = a_1 x + b_2 x^2 + \cdots$$

is regular and has modulus at most 1 for $|x| \leq 1$. But the function

$$f_1(x) = \frac{f(x)}{x(1 - f(x))}$$

is also regular for $|x| \leq 1$, and its value at $x = 0$ is a_1; by the maximum-modulus principle, its absolute value is at least as large at some point on $|x| = 1$. Since for $|x| = 1$,

$$|f_1(x)| = \left| \frac{f(x)}{1 - f(x)} \right|,$$

it follows that

$$|a_1| \leq 1. \tag{5}$$

The theorem will therefore be proved if we show that each of the functions

$$F_n(x) = a_n x + a_{2n} x^2 + \cdots$$

fulfills the same hypotheses as $f(x)$ itself. This depends on the fact that if $\eta = e^{2\pi i/n}$, then

$$\sum_{l=0}^{n-1} \eta^{lk} = \begin{cases} n & \text{if } n|k, \\ (\eta^{kn} - 1)/(\eta^k - 1) = 0 & \text{if } n \nmid k. \end{cases}$$

We have

$$\sum_{l=0}^{n-1} f(\eta^l x) = \sum_{l=0}^{n-1} \sum_{k=1}^{\infty} a_k \eta^{kl} x^k = \sum_{k=1}^{\infty} a_k x^k \sum_{l=0}^{n-1} \eta^{kl}$$

$$= n \sum_{n|k} a_k x^k = n F_n(x^n),$$

so that $F_n(x)$ is regular for $|x| \leq 1$, and for such x,

$$\operatorname{Re} F_n(x^n) = \frac{1}{n} \sum_{l=0}^{n-1} \operatorname{Re} f(\eta^l x) \leq \frac{1}{n} \sum_{l=0}^{n-1} \frac{1}{2} = \frac{1}{2}.$$

THEOREM 7–6. *Let R be positive, and suppose that*

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

is regular and $\operatorname{Re} f(x) \leq M$ *for* $|x - x_0| \leq R$. *Then, for* $n \geq 1$,

$$|a_n| \leq \frac{2}{R^n} (M - \operatorname{Re} a_0).$$

Proof: If $\operatorname{Re} a_0 = M$, then $a_n = 0$ for $n \geq 1$, by the maximum-modulus principle.

If $\operatorname{Re} a_0 < M$, put

$$g(x) = \frac{f(x_0 + Rx) - a_0}{2(M - \operatorname{Re} a_0)}.$$

Then g is regular for $|x| \leq 1$, $g(0) = 0$, and

$$\operatorname{Re} g(x) = \frac{\operatorname{Re} f(x_0 + Rx) - \operatorname{Re} a_0}{2(M - \operatorname{Re} a_0)} \leq \frac{M - \operatorname{Re} a_0}{2(M - \operatorname{Re} a_0)} = \frac{1}{2}.$$

Hence g satisfies the hypotheses of Theorem 7–5, so that

$$\left| \frac{a_n R^n}{2(M - \operatorname{Re} a_0)} \right| \leq 1,$$

and the theorem follows.

THEOREM 7–7. *If f satisfies the hypotheses of Theorem 7–6, and* $0 < r < R$, *then for* $|x - x_0| \leq r$,

$$|f(x)| \leq |a_0| + \frac{2r}{R - r} (|M| + |a_0|)$$

and
$$|f'(x)| \leq \frac{2R}{(R - r)^2} (|M| + |a_0|).$$

Proof: We have

$$|f(x)| \leq |a_0| + \sum_{n=1}^{\infty} |a_n| r^n \leq |a_0| + 2(|M| + |a_0|) \sum_{n=1}^{\infty} \left(\frac{r}{R} \right)^n$$

$$= |a_0| + \frac{2r}{R - r} (|M| + |a_0|),$$

and

$$|f'(x)| \le \sum_{n=1}^{\infty} |a_n| n r^{n-1} \le \frac{2(|M| + |a_0|)}{R} \sum_{n=1}^{\infty} n \left(\frac{r}{R}\right)^{n-1}$$

$$= \frac{2R}{(R-r)^2} (|M| + |a_0|).$$

THEOREM 7-8. *Let r be positive and M real, and suppose that $f(s_0) \ne 0$ and that, for $|s - s_0| \le r$, $f(s)$ is regular and*

$$\left| \frac{f(s)}{f(s_0)} \right| < e^M.$$

Suppose also that $f(s) \ne 0$ in the semicircular region $|s - s_0| \le r$, Re $s >$ Re s_0. Then

$$- \operatorname{Re} \frac{f'}{f}(s_0) \le \frac{4M}{r},$$

and if there is a zero ρ of f on the open line segment between $s_0 - r/2$ and s_0, then

$$- \operatorname{Re} \frac{f'}{f}(s_0) \le \frac{4M}{r} - \frac{1}{s_0 - \rho}.$$

Proof: There is clearly no loss in generality in supposing that $f(s_0) = 1$ and $s_0 = 0$. In this case, the hypotheses can be listed as follows.

(1) For $|s| \le r$, $f(s)$ is regular and $|f(s)| < e^M$, where $M > 0$.

(2) $f(0) = 1$.

(3) $f(s) \ne 0$ for $|s| \le r$, $\sigma > 0$.

We look for an upper bound for $- \operatorname{Re} f'(0)$.

If ρ runs through the zeros of f in the circle $|s| \le r/2$, then the function

$$g(s) = \frac{f(s)}{\prod_{\rho} (1 - s/\rho)}$$

is regular for $|s| \le r$. On the circle $|s| = r$, we have

$$\left| 1 - \frac{s}{\rho} \right| \ge \left| \frac{s}{\rho} \right| - 1 \ge 1,$$

so that here $|g(s)| \le |f(s)| < e^M$. By the maximum-modulus principle,

$$|g(s)| < e^M, \qquad \text{for } |s| \leq \frac{r}{2}.$$

Since $g(s) \neq 0$ for $|s| \leq r/2$, and $g(0) = 1$, we can write

$$g(s) = e^{G(s)}, \qquad \text{for } |s| \leq \frac{r}{2},$$

where G is regular and $\operatorname{Re} G(s) < M$, $G(0) = 0$. By Theorem 7–6, with $r/2$ instead of R,

$$|G'(0)| < \frac{2}{r/2} M = \frac{4M}{r}.$$

But

$$\frac{g'}{g}(s) = \frac{f'}{f}(s) - \sum_\rho \frac{-1/\rho}{1 - s/\rho} = \frac{f'}{f}(s) + \sum_\rho \frac{1}{\rho - s},$$

so that

$$\left| f'(0) + \sum_\rho \frac{1}{\rho} \right| = \left| \frac{g'}{g}(0) \right| = |G'(0)| \leq \frac{4M}{r},$$

$$-\operatorname{Re} f'(0) - \sum_\rho \frac{1}{\operatorname{Re} \rho} \leq \frac{4M}{r},$$

$$-\operatorname{Re} f'(0) \leq \frac{4M}{r} + \sum_\rho \operatorname{Re} \frac{1}{\rho}.$$

Since we have supposed that all zeros ρ have nonpositive real parts, the theorem follows.

If f is regular on the vertical line $\sigma_0 + ti$, and if

$$\lim_{\substack{a \to \infty \\ b \to \infty}} \int_{\sigma_0 - ai}^{\sigma_0 + bi} f(s)\, ds = \lim_{\substack{a \to \infty \\ b \to \infty}} \int_{-a}^{b} f(\sigma_0 + ti) i\, dt$$

exists, then we abbreviate this limit to

$$\int_{(\sigma_0)} f(s)\, ds.$$

THEOREM 7–9. *We have*

$$\frac{1}{2\pi i} \int_{(2)} \frac{y^s}{s^2}\, ds = \begin{cases} 0 & \text{for } 0 < y < 1, \\ \log y & \text{for } y \geq 1. \end{cases}$$

Proof: The integral converges, because

$$\left|\frac{y^s}{s^2}\right| = \frac{y^2}{4 + t^2}.$$

First suppose that $0 < y < 1$. Then in the region bounded by C_1 and C_2 (see Fig. 7-1), the integrand is regular, so that by Cauchy's theorem,

FIGURE 7-1

$$\int_{C_1} \frac{y^s}{s^2}\, ds + \int_{C_2} \frac{y^s}{s^2}\, ds = 0.$$

But along C_2, which is of length πa, we have

$$\left|\frac{y^s}{s^2}\right| < \frac{1}{a^2},$$

so that

$$\left|\int_{C_2} \frac{y^s}{s^2}\, ds\right| < \frac{\pi a}{a^2} = \frac{\pi}{a}.$$

Hence, as $a \to \infty$,

$$\int_{C_2} \frac{y^s}{s^2}\, ds \longrightarrow 0, \qquad \int_{C_1} \frac{y^s}{s^2}\, ds \longrightarrow \int_{(2)} \frac{y^s}{s^2}\, ds,$$

and the result follows.

Now suppose that $y \geq 1$, and that $a > 2$. Then the pole $s = 0$ of the integrand lies in the region bounded by C_1 and C_3, and since

$$\frac{y^s}{s^2} = \frac{1 + s \log y + (s^2 \log^2 y)/2 + \cdots}{s^2} = \frac{1}{s^2} + \frac{\log y}{s} + \cdots,$$

we have by the residue theorem that

$$\frac{1}{2\pi i} \int_{C_1} \frac{y^s}{s^2}\, ds + \frac{1}{2\pi i} \int_{C_3} \frac{y^s}{s^2}\, ds = \log y.$$

But along C_3, which is of length πa, we have

$$\left|\frac{y^s}{s^2}\right| = \frac{y^\sigma}{|s|^2} \leq \frac{y^2}{(a - 2)^2},$$
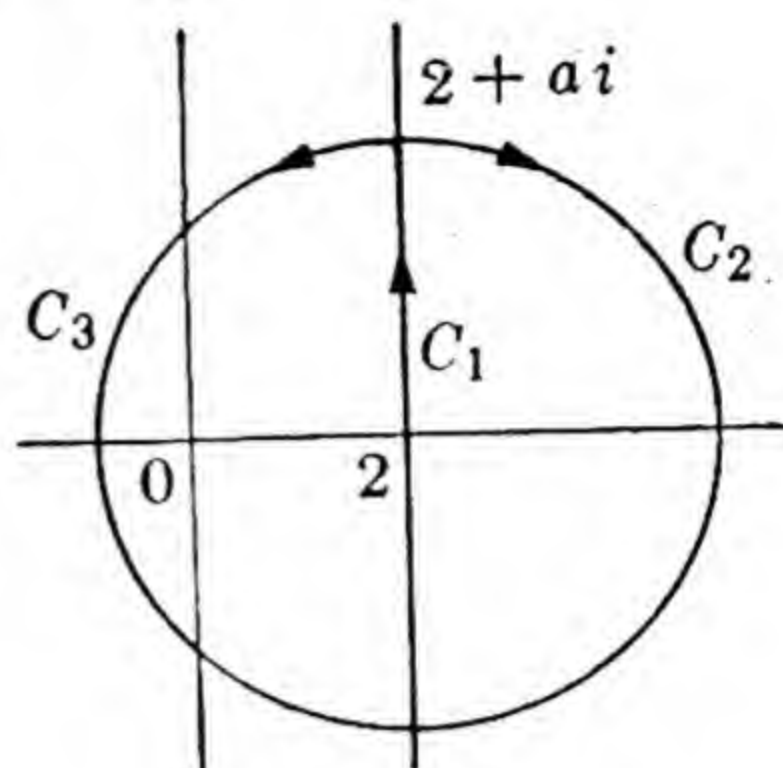
so that, for $a > 4$,

$$\left| \int_{C_3} \frac{y^s}{s^2}\, ds \right| < \frac{\pi a y^2}{(a-2)^2} < \frac{4\pi y^2}{a}.$$

Hence, as $a \to \infty$,

$$\int_{C_3} \frac{y^s}{s^2}\, ds \longrightarrow 0, \qquad \int_{C_1} \frac{y^s}{s^2}\, ds \longrightarrow \int_{(2)} \frac{y^s}{s^2}\, ds,$$

and the result follows.

7–3 The Prime Number Theorem. It will be necessary in what follows to know something about the location of the zeros of the ζ-function. For $|t|$ large, this information is supplied by Theorem 7–13; for $|t|$ small, we use only the fact that $\zeta(s)$ *does not vanish for $\sigma \geq 1$*. Historically, this was the first nontrivial result obtained concerning the zeros of the ζ-function. (A trivial fact is that $\zeta(s) \neq 0$ for $\sigma > 1$, which follows immediately from the product representation

$$\zeta(s) = \prod_p (1 - p^{-s})^{-1},$$

valid for $\sigma > 1$.) The proof below that $\zeta(1 + ti) \neq 0$ is due to de la Vallée Poussin; it may have been suggested by the following considerations.

For $\sigma > 1$ we have

$$\log \zeta(s) = \sum_{m,p} \frac{1}{m p^{ms}} = \sum_p \frac{1}{p^s} + f(s),$$

and f is easily seen to be regular for $\sigma > \frac{1}{2}$. Since ζ has a pole at $s = 1$, with residue 1, it follows that as $\sigma \to 1^+$,

$$\sum_p \frac{1}{p^\sigma} \sim \log \frac{1}{\sigma - 1}. \tag{6}$$

We now reason heuristically. If $1 + t_0 i$ is a zero of ζ, and we put $s = \sigma + t_0 i$, then as $\sigma \to 1^+$,

$$\log |\zeta(s)| \sim \log (\sigma - 1)$$

and

$$\mathrm{Re}\, \log \zeta(s) - \mathrm{Re}\, f(s) = \log |\zeta(s)| - \mathrm{Re}\, f(s)$$

$$= \sum_p \frac{\cos (t_0 \log p)}{p^\sigma} \sim \log (\sigma - 1).$$

Comparing this with (6) we see that for most p, $\cos (t_0 \log p)$ must

be close to -1. But then $\cos(2t_0 \log p)$ must usually be nearly 1, and

$$\sum_p \frac{\cos(2t_0 \log p)}{p^\sigma} \sim \log \frac{1}{\sigma - 1}.$$

But this requires that ζ have a pole at $1 + 2t_0 i$, which is not the case. To make this argument rigorous, note that for all real θ,

$$3 + 4\cos\theta + \cos 2\theta = 2(1 + \cos\theta)^2 \geq 0.$$

Hence for $\sigma > 1$,

$$\log|\zeta^3(\sigma)\zeta^4(\sigma + t_0 i)\zeta(\sigma + 2t_0 i)|$$

$$= 3\log|\zeta(\sigma)| + 4\log|\zeta(\sigma + t_0 i)| + \log|\zeta(\sigma + 2t_0 i)|$$

$$= 3\sum_{n,p}\frac{1}{np^{n\sigma}} + 4\sum_{n,p}\frac{\cos(t_0 n \log p)}{np^{n\sigma}} + \sum_{n,p}\frac{\cos(2t_0 n \log p)}{np^{n\sigma}}$$

$$= \sum_{n,p}\frac{3 + 4\cos(t_0 n \log p) + \cos(2t_0 n \log p)}{np^{n\sigma}}$$

$$\geq 0.$$

Thus

$$((\sigma - 1)\zeta(\sigma))^3 \left|\frac{\zeta(\sigma + t_0 i)}{\sigma - 1}\right|^4 |\zeta(\sigma + 2t_0 i)| \geq \frac{1}{\sigma - 1},$$

and if $1 + t_0 i$ were a zero of ζ, the left side in this inequality would remain bounded as $\sigma \to 1^+$, while the right side increases without limit.

We now use this technique, together with Theorem 7–8, to show that $\zeta(s)$ does not vanish at any point too close to the line $\sigma = 1$ and sufficiently far from the real axis.

THEOREM 7–10.   *For $\sigma > 1$,*

$$\mathrm{Re}\left(-3\frac{\zeta'}{\zeta}(\sigma) - 4\frac{\zeta'}{\zeta}(\sigma + ti) - \frac{\zeta'}{\zeta}(\sigma + 2ti)\right) \geq 0.$$

Proof: Differentiating the relation

$$\log\zeta(s) = \sum_{m,p}\frac{1}{mp^{ms}},$$

we obtain

$$\frac{\zeta'}{\zeta}(s) = -\sum_{m,p}\frac{\log p}{p^{ms}} = -\sum_{n=1}^{\infty}\frac{\Lambda(n)}{n^s}, \tag{7}$$

where

$$\Lambda(n) = \begin{cases} \log p & \text{if } n = p^m, \text{ for any } m > 0 \text{ and prime } p, \\ 0 & \text{otherwise.} \end{cases}$$

The termwise differentiation is justified because the series for $\log \zeta(s)$ converges uniformly in any region to the right of $\sigma = 1$. Hence

$$\text{Re} \left(-3 \frac{\zeta'}{\zeta}(\sigma) - 4 \frac{\zeta'}{\zeta}(\sigma + ti) - \frac{\zeta'}{\zeta}(\sigma + 2ti) \right)$$

$$= \text{Re} \sum_{n=1}^{\infty} \frac{(3 + 4n^{-ti} + n^{-2ti})\Lambda(n)}{n^{\sigma}}$$

$$= \sum_{n=1}^{\infty} \frac{(3 + 4\cos(t \log n) + \cos(2t \log n))\Lambda(n)}{n^{\sigma}}$$

$$\geq 0.$$

THEOREM 7-11. (a) *For $\sigma \geq \frac{1}{2}$ and $t > c$, we have $|\zeta(s)| < t$.*
(b) *For $t \geq 8$ and $\sigma \geq 1 - (\log t)^{-1}$, we have $|\zeta(s)| < c \log t$.*

Proof: For $\sigma > 2$ and $t \geq 8$,

$$|\zeta(s)| < \sum_{n=1}^{\infty} \frac{1}{n^2} < 2 < \begin{cases} t, \\ \log t. \end{cases}$$

For $\sigma \leq 2$, both inequalities of the theorem follow from Theorem 7-4.

THEOREM 7-12. For $\sigma > 1$,

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \frac{\mu(n)}{n^s},$$

where μ is the Möbius function.

Proof: This follows immediately from Theorem 6-12 for s real and greater than 1; by analytic continuation, it is correct for $\sigma > 1$.

THEOREM 7-13. *There are constants $c_1 > 8$ and $c_2 > 0$ such that $\zeta(s) \neq 0$ for*

$$t > c_1 \quad \text{and} \quad \sigma > 1 - \frac{c_2}{\log t}$$

Proof: In accordance with Theorem 7-11 (a), choose that

$$|\zeta(s)| < t, \quad \text{for } \sigma \geq \tfrac{1}{2}, \quad t > c_3.$$

$$\sigma = 1 - \frac{c_2}{\log t}$$

$$\sigma = 1 + \frac{c_4}{\log t}$$

$$\beta + i\gamma$$

$$\sigma = \frac{3}{4} + \frac{c_4}{\log t}$$
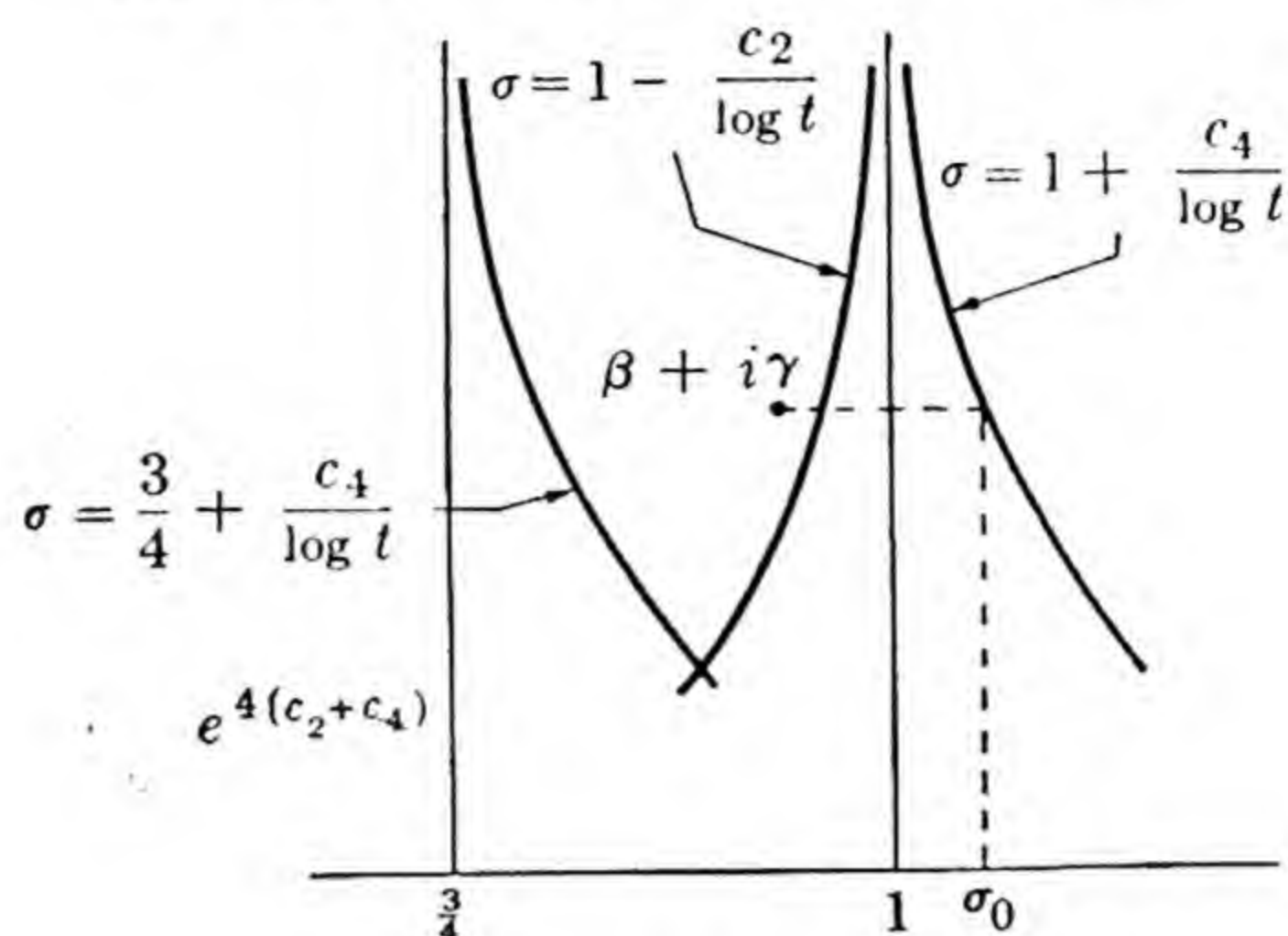
$$e^{4(c_2+c_4)}$$

FIGURE 7-2

Inasmuch as

$$\frac{3}{4} + \frac{c_4}{\log x} < 1 - \frac{c_2}{\log x}, \qquad \text{for } x > e^{4(c_2+c_4)},$$

it suffices to show that any zero $\beta + \gamma i$ of ζ with γ sufficiently large (in particular, larger than 8) and for which

$$\beta > \frac{3}{4} + \frac{c_4}{\log \gamma},$$

is such that

$$\beta < 1 - \frac{c_2}{\log \gamma}.$$

Put

$$\sigma_0 = \sigma_0(\gamma) = 1 + \frac{c_4}{\log \gamma},$$

and suppose that $\beta + \gamma i$ is a zero of ζ for which $\gamma > e^{4(1+c_2+c_4)}$ and $\beta > \sigma_0 - \frac{1}{4}$. We shall apply Theorem 7-8, once with $s_0 = \sigma_0 + \gamma i$ and once with $s_0 = \sigma_0 + 2\gamma i$. In either case, since $\sigma_0 > 1$, we have that for $\gamma \geq c_3 + \frac{1}{2}$, the circle $|s - s_0| \leq \frac{1}{2}$ lies in the quadrant $\sigma \geq \frac{1}{2}, t \geq c_3$. Since $\gamma > e^{c_4}$, we have $\sigma_0 < 2$, and, by Theorem 7-12,

$$\frac{1}{\sigma_0} < 1 + \int_1^\infty \frac{du}{u^{\sigma_0}} = 1 + \frac{1}{\sigma_0 - 1} < \frac{2}{\sigma_0 - 1} = \frac{2}{c_4} \log \gamma.$$

Thus for each $\epsilon_1 > 0$ there is a c_5 such that for $\gamma > c_5 > c_3 + \frac{1}{2}$, the ineq.

$$\left| \frac{\zeta(s)}{\zeta(s_0)} \right| \leq \frac{2}{c_4} \left(2\gamma + \frac{1}{2} \right) \log \gamma < \gamma^{1+\epsilon_1}$$

holds at every point s of the circular disk $|s - s_0| \leq \frac{1}{2}$, since at every such point, $c_3 < t \leq 2\gamma + \frac{1}{2}$. If $\gamma \geq c_5$, we can now apply Theorem 7–8, with $r = \frac{1}{2}$, $f(s) = \zeta(s)$, $M = (1 + \epsilon_1) \log \gamma$. Using the first inequality of that theorem with $s_0 = \sigma_0 + 2\gamma i$, we obtain

$$-\operatorname{Re} \frac{\zeta'}{\zeta} (\sigma_0 + 2\gamma i) < 8(1 + \epsilon_1) \log \gamma; \tag{9}$$

using the second with $s_0 = \sigma_0 + \gamma i$ we have

$$-\operatorname{Re} \frac{\zeta'}{\zeta} (\sigma_0 + \gamma i) < 8(1 + \epsilon_1) \log \gamma - \frac{1}{\sigma_0 - \beta}, \tag{10}$$

since

$$\sigma_0 - r/2 = \sigma_0 - \tfrac{1}{4} < \beta \leq 1 < \sigma_0.$$

Finally, since $\sigma_0 \to 1^+$ as $t \to \infty$, we have from (6) that for $\epsilon_2 > 0$,

$$-\frac{\zeta'}{\zeta} (\sigma_0) < \frac{1 + \epsilon_2}{\sigma_0 - 1} = \frac{1 + \epsilon_2}{c_4} \log \gamma \tag{11}$$

for $\gamma > c_6$. Using the estimates (9), (10), and (11) in Theorem 7–10 gives

$$\frac{3(1 + \epsilon_2)}{c_4} \log \gamma + 4 \cdot 8(1 + \epsilon_1) \log \gamma - \frac{4}{\sigma_0 - \beta} + 8(1 + \epsilon_1) \log \gamma \geq 0.$$

This inequality can easily be simplified to

$$\sigma_0 - \beta > \frac{c_7}{\log \gamma},$$

where

$$c_7 = \frac{4c_4}{3(1 + \epsilon_2) + 40(1 + \epsilon_1)c_4},$$

and this gives

$$\beta < 1 - \frac{c_7 - c_4}{\log \gamma}.$$

It is clear that $c_7 > c_4$ if $\epsilon_1 < \frac{1}{3}$ and c_4 is sufficiently small, and we can then take $c_2 = c_7 - c_4$ and $c_1 = \max (c_5, c_6)$.

THEOREM 7–14. *If $0 < c_8 < c_2$, then*

$$|\log \zeta(s)| < \log^2 t \qquad \textit{for } t > c_9 \textit{ and } \sigma \geq 1 - \frac{c_8}{\log t}.$$

Proof: We use Theorem 7–7, with $s_0 = 2 + t_0 i$, for some $t_0 > 8$ to be determined. For t sufficiently large, the circular region

$$|s - s_0| \leq 1 + \frac{\frac{1}{2}(c_2 + c_8)}{\log t_0} \tag{12}$$

lies entirely in the region described in the preceding theorem, in which ζ has no zeros. Hence the function $\log \zeta(s)$ is regular in this disk, and by Theorem 7–11(b),

$$\text{Re} \log \zeta(s) = \log |\zeta(s)| < \log (c \log t)$$
$$< \log (c \log (t_0 + 2)) < c_{10} \log \log t_0.$$

Hence, by Theorem 7–7, we have that for s in the region (12),

$$|\log \zeta(s)| \leq |\zeta(s_0)| + \frac{2 \cdot 2(c_{10} \log \log t_0 + |\zeta(s_0)|)}{\dfrac{c_8 - c_2}{2} \cdot \dfrac{1}{\log t_0}}$$

$$\leq c + c \log t_0 \log \log t_0 < \log^2 t_0,$$

if t_0 is sufficiently large. This inequality holds on the radius extending toward the left from s_0, for every large t_0, and hence throughout a region $t \geq c_9$, $1 - c_8 (\log t)^{-1} \leq \sigma \leq 2$. Finally, $|\zeta(s)|$ and $|1/\zeta(s)|$ are bounded in the half-plane $\sigma > 2$, and $|\log \zeta(s)|$ is consequently smaller than $\log^2 t$ for t large and $\sigma > 2$.

THEOREM 7–15. *There is a constant $\alpha > 0$ such that as $x \to \infty$,*

$$\sum_{p \leq x} \log \frac{x}{p} = \int_c^1 \frac{x^s}{s^2} ds + O(xe^{-\alpha\sqrt{\log x}}),$$

for some c with $0 < c < 1$.

Proof: Using Theorem 7–9, we have

$$\frac{1}{2\pi i} \int_{(2)} \frac{x^s}{s^2} \log \zeta(s) \, ds = \frac{1}{2\pi i} \int_{(2)} \frac{1}{s^2} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{\log n} \left(\frac{x}{n}\right)^s ds$$

$$= \frac{1}{2\pi i} \sum_{n=1}^{\infty} \frac{\Lambda(n)}{\log n} \int_{(2)} \frac{(x/n)^s}{s^2} \, ds = \sum_{n \leq x} \frac{\Lambda(n)}{\log n} \log \frac{x}{n} = \sum_{\substack{m,p \\ p^m \leq x}} \frac{1}{m} \log \frac{x}{p^m}$$

$$= \sum_{p \leq x} \log \frac{x}{p} + \sum_{\substack{m,p \\ m \geq 2 \\ p^m \leq x}} \frac{1}{m} \log \frac{x}{p^m}.$$

FIGURE 7–3

As noted earlier, the number of terms in the last sum is

$$\pi(x^{\frac{1}{2}}) + \pi(x^{\frac{1}{3}}) + \cdots < x^{\frac{1}{2}} + x^{\frac{1}{3}} + \cdots + x^{1/u} < ux^{\frac{1}{2}},$$

where u is the smallest integer such that $x^{1/u} < 2$. Thus

$$\pi(x^{\frac{1}{2}}) + \pi(x^{\frac{1}{3}}) + \cdots = O(x^{\frac{1}{2}} \log x),$$

so that

$$2\pi i \sum_{p \le x} \log \frac{x}{p} = \int_{(2)} \frac{x^s}{s^2} \log \zeta(s)\, ds + O(xe^{-\sqrt{\log x}}),$$

since

$$\sum_{\substack{m \ge 2 \\ p^m \le x}} \frac{1}{m} \log \frac{x}{p^m} \le \sum_{\substack{m \ge 2 \\ p^m \le x}} \log x = O(\sqrt{x} \log^2 x) = O(xe^{-\sqrt{\log x}}).$$

We now cut the complex plane along the real axis, extending the cut from $s = 1$ to the left, and examine the function $\log \zeta(s)$ in the cut plane. If $\bar{z}$ is the complex conjugate of z, then $\zeta(\bar{s}) = \overline{\zeta(s)}$ and $\log \bar{z} = \overline{\log z}$, so that $\log \zeta(\bar{s}) = \overline{\log \zeta(s)}$. Hence, by Theorem 7–13, $\zeta(s) \neq 0$ for $|t| > c_9 > c_1$ and $\sigma \ge 1 - c_8 (\log |t|)^{-1}$. Moreover, since $\zeta(s)$ does not vanish on the line $\sigma = 1$, and since its zeros have no

finite limit point in any half-plane $\sigma \geq \sigma_0 > 0$ (since $\zeta(s)$ is regular there), there is a constant $c_{11} > 0$ such that $\zeta(s) \neq 0$ in the rectangle

$$1 - c_{11} \leq \sigma \leq 1, \qquad |t| \leq c_9.$$

Finally, $\zeta(s) \neq 0$ for $1 \leq \sigma \leq 2$, and the only singularity of the function in the half-plane $\sigma > 0$ is at $s = 1$. Consequently, for arbitrary $u > c_9$, $\log \zeta(s)$ is a single-valued analytic function in the region Q shown in Fig. 7–3, bounded by the arcs $\Gamma_1, \Gamma_2, \ldots, \Gamma_6, \Gamma_7,$ $\bar{\Gamma}_6, \ldots, \bar{\Gamma}_2, \bar{\Gamma}_1$. Hence if we denote by Γ the complete boundary of this region (so that we might write symbolically $\Gamma = \Gamma_1 + \Gamma_2 + \cdots$ $+ \bar{\Gamma}_1$), we have by Cauchy's theorem that

$$\int_\Gamma \frac{x^s}{s^2} \log \zeta(s) \, ds = 0.$$

It follows that, if the integrals are taken in the positive direction,

$$\int_{(2)} \frac{x^s}{s^2} \log \zeta(s) \, ds$$

$$= \left(\int_{2-\infty i}^{2-ui} + \int_{\bar{\Gamma}_1} + \int_{\Gamma_1} + \int_{2+ui}^{2+\infty i} \right) \frac{x^s}{s^2} \log \zeta(s) \, ds$$

$$= \left(\int_{2-\infty i}^{2-ui} - \int_{\Gamma_2 + \cdots + \Gamma_6 + \Gamma_7 + \bar{\Gamma}_6 + \cdots + \bar{\Gamma}_2} + \int_{2+ui}^{2+\infty i} \right) \frac{x^s}{s^2} \log \zeta(s) \, ds.$$

We shall show that all the other integrals are small in comparison with those along Γ_6 and $\bar{\Gamma}_6$, if u is sufficiently large. For brevity, put

$$\psi(x, s) = \frac{x^s}{s^2} \log \zeta(s).$$

By Theorem 7–14, we have that for $u > u_0(\epsilon)$,

$$\left| \int_{2+ui}^{2+\infty i} \psi(x, s) \, ds \right| \leq \int_{2+ui}^{2+\infty i} \frac{x^2}{|s|^2} |\log \zeta(s)| |ds|$$

$$\leq x^2 \int_u^\infty \frac{\log^2 t}{t^2} \, dt \leq x^2 \int_u^\infty \frac{dt}{t^{2-\epsilon}} < \frac{cx^2}{u^{1-\epsilon}},$$

so that

$$\lim_{u \to \infty} \int_{2+ui}^{2+\infty i} \psi(x, s) \, ds = 0.$$

The same estimate applies to the integral from $2 - \infty i$ to $2 - ui$.

Since the length of Γ_2 is less than 2, and the integrand is again smaller than $x^2 \log^2 u/u^2$ for u large, we have

$$\lim_{u \to \infty} \int_{\Gamma_2} \psi(x, s)\, ds = 0,$$

and similar considerations give

$$\lim_{u \to \infty} \int_{\bar{\Gamma}_2} \psi(x, s)\, ds = 0.$$

Along Γ_3 we have $s = 1 - c_8 (\log t)^{-1} + ti$, so that

$$\left| \int_{\Gamma_3} \psi(x, s)\, ds \right| \leq \int_{c_9}^{u} \frac{x^{1-c_8(\log t)^{-1}}}{t^2} \log^2 t \left| \frac{c_8}{t \log^2 t} + i \right| dt.$$

Now suppose that x, and then u, are chosen large enough that

$$c_9 < e^{\sqrt{2c_8 \log x}} < u.$$

Then

$$\int_{\Gamma_3} \psi(x, s)\, ds$$

$$= O\left(\int_{c_9}^{e^{\sqrt{2c_8 \log x}}} x \cdot x^{-c_8(2c_8 \log x)^{-\frac{1}{2}}} \frac{\log^2 t}{t^2} dt + x \int_{e^{\sqrt{2c_8 \log x}}}^{u} \frac{\log^2 t}{t^{\frac{1}{2}} \cdot t^{\frac{3}{2}}} dt \right)$$

$$= O\left(xe^{-\sqrt{\frac{1}{2}c_8 \log x}} \int_{c_9}^{\infty} \frac{\log^2 t}{t^2} dt + \frac{x}{e^{\sqrt{\frac{1}{2}c_8 \log x}}} \cdot \int_{e^{\sqrt{2c_8 \log x}}}^{u} \frac{\log^2 t\, dt}{t^{\frac{3}{2}}} \right)$$

$$= O(xe^{-\alpha\sqrt{\log x}}),$$

where $\alpha = \sqrt{c_8/2}$.

By symmetry,

$$\int_{\bar{\Gamma}_3} \psi(x, s)\, ds = O(xe^{-\alpha\sqrt{\log x}}).$$

The paths Γ_4, Γ_5, $\bar{\Gamma}_4$, and $\bar{\Gamma}_5$ are of fixed lengths, and on them

$$\psi(x, s) = O(x^{1-c_{11}}) = o(xe^{-\alpha\sqrt{\log x}}),$$

so that the same estimate holds for the integrals themselves.

Γ_7 is described by the relations $s = 1 + \delta e^{i\theta}$, $|\theta| \leq \pi$, where $\delta > 0$. Since $(s-1)\zeta(s) \to 1$ as $s \to 1$, we have

$$\text{Re} \log \zeta(s) = \log |\zeta(s)| \sim -\log |s-1| = -\log \delta,$$

$$\text{Im} \log \zeta(s) = \arg \zeta(s) = O(1)$$

as $\delta \to 0^+$. Hence

$$\int_{\Gamma_7} \psi(x, s)\, ds = O\left(2\pi\delta \frac{x^{1+\delta}}{(1-\delta)^2} \log \delta\right) = o(1).$$

Combining all these results, we can take the limit as $u \to \infty$ and $\delta \to 0^+$ and obtain

$$2\pi i \sum_{p \leq x} \log \frac{x}{p} = \int_{1-c_{11}}^{1} \psi(x, s)\, ds + \int_{1}^{1-c_{11}} \psi(x, s)\, ds + o(xe^{-\alpha\sqrt{\log x}}),$$

where the first integral is along the upper edge, and the second along the lower edge, of the cut. We know that $(1-s)\zeta(s) = R(s)$ is regular in the region $\sigma > 0$, and that it has no zeros in the region $\sigma > 1 - c_{11}$, $|t| < c_9$. Hence the function

$$\log ((s-1)\zeta(s)) = \log (s-1) + \log \zeta(s)$$

is single-valued in this region; since $\log (s-1)$ has, on the upper and lower edges of the cut, values which differ by $2\pi i$, the same is true of $\log \zeta(s)$, if the difference is taken in reverse order. Hence if s^+ indicates the upper edge of the cut, and s^- the lower edge, then

$$\int_{1-c_{11}}^{1} \psi(x, s^+)\, ds^+ + \int_{1}^{1-c_{11}} \psi(x, s^-)\, ds^-$$

$$= \int_{1-c_{11}}^{1} \frac{x^{s^+}}{(s^+)^2} \log \zeta(s^+)\, ds^+ - \int_{1-c_{11}}^{1} \frac{x^{s^+}}{(s^+)^2} (\log \zeta(s^+) - 2\pi i)\, ds^+$$

$$= 2\pi i \int_{1-c_{11}}^{1} \frac{x^s}{s^2}\, ds,$$

and

$$\sum_{p \leq x} \log \frac{x}{p} = \int_{1-c_{11}}^{1} \frac{x^s}{s^2}\, ds + O(xe^{-\alpha\sqrt{\log x}}). \tag{13}$$

The theorem is proved.

THEOREM 7-16.   *As $x \to \infty$,*

$$\pi(x) = \int_2^x \frac{du}{\log u} + O(xe^{-\frac{1}{2}\alpha\sqrt{\log x}}).$$

Proof: Replace $1 - c_{11}$ by C in (13), and put

$$\delta = \delta(x) = e^{-\frac{1}{2}\alpha\sqrt{\log x}}.$$

Then since $\log (1 + \delta) \sim \delta$ as $x \to \infty$, we have

$$\sum_{p \leq x(1+\delta)} \log \frac{x(1 + \delta)}{p} - \sum_{p \leq x} \log \frac{x}{p}$$

$$= \sum_{p \leq x} \log (1 + \delta) + \sum_{x < p \leq x(1+\delta)} \log \frac{x(1 + \delta)}{p}$$

$$= \log (1 + \delta)\pi(x) + O(\log (1 + \delta) \cdot \delta x)$$

$$= \int_C^1 \frac{x^s}{s^2} ((1 + \delta)^s - 1)\, ds + O(xe^{-\alpha\sqrt{\log x}}),$$

so that

$$\pi(x) = \frac{1}{\log (1 + \delta)} \int_C^1 \frac{(1 + \delta)^s - 1}{s^2} x^s\, ds + O(\delta x) + O\left(\frac{xe^{-\alpha\sqrt{\log x}}}{\delta}\right).$$

Now

$$(1 + \delta)^s - 1 = s\delta + \frac{s(s - 1)}{2!} (1 + \vartheta\delta)^{s-2}\delta^2,$$

where $0 < \vartheta < 1$, so that for $0 < s < 1$,

$$|(1 + \delta)^s - 1 - s\delta| \leq \frac{s|s - 1|}{2} \delta^2 < \delta^2.$$

Thus, making the change of variable $x^s = u$, we obtain

$$\int_C^1 \frac{(1 + \delta)^s - 1}{s^2} x^s\, ds = \delta \int_C^1 \frac{x^s}{s}\, ds + O\left(\delta^2 \int_C^1 \frac{x^s}{s^2}\, ds\right)$$

$$= \delta \int_{x^C}^x \frac{du}{\log u} + O\left(\delta^2 \int_C^1 x^s\, ds\right)$$

$$= \delta \int_2^x \frac{du}{\log u} + O(\delta^2 x).$$

Finally,

$$\pi(x) = \frac{\delta}{\log(1+\delta)} \int_2^x \frac{du}{\log u} + O(\delta x) + O\left(\frac{xe^{-\alpha\sqrt{\log x}}}{\delta}\right)$$

$$= (1 + O(\delta)) \int_2^x \frac{du}{\log u} + O(\delta x) + O(xe^{-\frac{1}{2}\alpha\sqrt{\log x}})$$

$$= \int_2^x \frac{du}{\log u} + O(\delta x) = \int_2^x \frac{du}{\log u} + O(xe^{-\frac{1}{2}\alpha\sqrt{\log x}}).$$

The Prime Number Theorem is a very weak consequence of Theorem 7-16, since

$$\int_2^x \frac{du}{\log u} = \frac{u}{\log u}\Big]_2^x + \int_2^x \frac{du}{\log^2 u} \sim \frac{x}{\log x}$$

and

$$xe^{-c\sqrt{\log x}} = o\left(\frac{x}{\log x}\right)$$

for every $c > 0$. In fact, we see by repeated integration by parts that the relation

$$\pi(x) = \frac{x}{\log x} + \frac{2!x}{\log^2 x} + \frac{3!x}{\log^3 x} + \cdots + \frac{m!x}{\log^m x} + o\left(\frac{x}{\log^m x}\right)$$

holds for every positive integer m.

The coefficient α occurring in the remainder term in Theorem 7-16 can easily be bounded explicitly; it can be shown for example that $\alpha = \frac{1}{45}$ is an allowable value, by choosing $c_2 = \frac{1}{1000}$, $c_4 = \frac{1}{200}$, $c_8 = \frac{1}{1002}$, $\epsilon_1 = \frac{1}{10}$, $\epsilon_2 = \frac{3}{100}$. However, no result of this type is as good as the known result that

$$\pi(x) = \int_2^x \frac{du}{\log u} + O(xe^{-c\sqrt{\log x \log\log x}}).$$

In a variant of the proof given here, the factor $\log \zeta(s)$ in the integrand is replaced by $\zeta'(s)/\zeta(s)$. The logarithmic singularity at $s = 1$ is then replaced by a simple pole, which makes the analysis somewhat less complicated. On the other hand this gives an estimate not of

$$\pi(x) = \sum_{p \le x} 1,$$

but of

$$\psi(x) = \sum_{p \le x} \log p \left[\frac{\log x}{\log p}\right],$$

and an additional step is needed to obtain the final result.

7–4 Extension to primes in an arithmetic progression. For relatively prime integers k and l, let $\pi(x; k, l)$ be the number of primes $p \equiv l \pmod{k}$ which do not exceed x. For given k, there are $\varphi(k) = h$ choices of l which are distinct modulo k, so that if the primes are more or less evenly dispersed among the various progressions, it is to be expected that

$$\pi(x; k, l) \sim \frac{1}{h} \frac{x}{\log x}.$$

It is the object of this section to show that this is the case, and in fact to obtain an estimate for $\pi(x; k, l)$ similar to that given in Theorem 7–16 for $\pi(x) = \pi(x; 1, 0)$. Several proofs will not be given in full detail since they are similar to those of the preceding sections. As in the preceding chapter, we isolate the primes lying in a given arithmetic progression by use of characters and L-functions. The L-functions are in turn simple combinations of Hurwitz ζ-functions, as the following theorem shows.

THEOREM 7–17. *For $\sigma > 1$,*

$$L(s, \chi) = \frac{1}{k^s} \sum_{a=1}^{k} \chi(a) \zeta\left(s, \frac{a}{k}\right). \tag{14}$$

Proof: Since χ is periodic of period k,

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

$$= \sum_{a=1}^{k} \chi(a) \sum_{m=0}^{\infty} \frac{1}{(km + a)^s}$$

$$= \frac{1}{k^s} \sum_{a=1}^{k} \chi(a) \zeta\left(s, \frac{a}{k}\right).$$

The domain of validity of (14) can be extended somewhat. If we put

$$E(\chi) = \begin{cases} 1 & \text{if } \chi = \chi_0, \\ 0 & \text{if } \chi \neq \chi_0, \end{cases}$$

then the first equation of Theorem 6–6 becomes

$$\sum_{a=1}^{k} \chi(a) = E(\chi)h.$$

Hence, for $\sigma > 1$,

$$L(s, \chi) - \frac{E(\chi)h}{k} \cdot \frac{1}{s-1}$$
$$= \frac{E(\chi)h}{s-1} \left(\frac{1}{k^s} - \frac{1}{k} \right) + \frac{1}{k^s} \sum_{a=1}^{k} \chi(a) \left\{ \zeta\left(s, \frac{a}{k}\right) - \frac{1}{s-1} \right\} \cdot$$

By Theorem 7-3, each summand on the right is regular for $\sigma > 0$, and

$$\frac{1}{s-1} \left(\frac{1}{k^s} - \frac{1}{k} \right) = \frac{k^{1-s} - 1}{k(s-1)}$$

is an integral function. By analytic continuation, we have

THEOREM 7-18. *The relation* (14) *holds for* $\sigma > 0$, *except at* $s = 1$. *Moreover,*

$$\lim_{s \to 1} (s - 1)L(s, \chi) = \frac{hE(\chi)}{k} \cdot \tag{15}$$

Hence $L(s, \chi)$ *is regular for* $\sigma > 0$, *except that* $L(s, \chi_0)$ *has a simple pole at* $s = 1$.

For $\sigma > 2$ and $t \geq 8$,

$$|L(s, \chi)| < \sum_{n=1}^{\infty} \frac{1}{n^2} < 2 < \begin{cases} t, \\ \log t, \end{cases}$$

while for $\sigma > 0$ and $t > 0$,

$$|L(s, \chi)| \leq \sum_{a=1}^{k} \left| \zeta\left(s, \frac{a}{k}\right) \right|,$$

so that Theorem 7-4 yields

THEOREM 7-19. (a) *For* $\sigma \geq \frac{1}{2}$ *and* $t > c_{12}(k)$, *we have* $|L(s, \chi)| < t$.
(b) *For* $t \geq 8$ *and* $\sigma > 1 - (\log t)^{-1}$, *we have* $|L(s, \chi)| < c_{13}(k) \log t$.

The proof of the nonvanishing of $\zeta(s)$ on $\sigma = 1$ can be generalized in a simple way.

THEOREM 7-20. $L(s, \chi)$ *does not vanish on the line* $\sigma = 1$.

Proof: For $\sigma > 1$,

$$L(s, \chi) = \prod_{p} (1 - \chi(p)p^{-s})^{-1},$$

so that we can choose

$$\log L(s, \chi) = \sum_{m,p} \frac{\chi(p^m)}{mp^{ms}}.$$

Hence

$$\log |L^3(\sigma, \chi_0)L^4(\sigma + ti, \chi)L(\sigma + 2ti, \chi^2)|$$
$$= 3 \log |L(\sigma, \chi_0)| + 4 \log |L(\sigma + ti, \chi)| + \log |L(\sigma + 2ti, \chi^2)|$$
$$= 3 \log L(\sigma, \chi_0) + 4 \operatorname{Re} \log L(\sigma + ti, \chi) + \operatorname{Re} \log L(\sigma + 2ti, \chi^2)$$
$$= \sum_{m,p} \left(\frac{3\chi_0(p^m)}{mp^{m\sigma}} + \operatorname{Re} \frac{4\chi(p^m)}{mp^{m(\sigma+ti)}} + \operatorname{Re} \frac{\chi^2(p^m)}{mp^{m(\sigma+2ti)}} \right)$$
$$= \sum_{\substack{m,p \\ p \nmid k}} \frac{3 + 4 \cos (\eta(p^m) - t \log p^m) + \cos 2(\eta(p^m) - t \log p^m)}{mp^{m\sigma}}$$
$$\geq 0,$$

where $\chi(p^m) = e^{i\eta(p^m)}$. Thus

$$((\sigma - 1)L(\sigma, \chi_0))^3 \left| \frac{L(\sigma + ti, \chi)}{\sigma - 1} \right|^4 |L(\sigma + 2ti, \chi^2)| \geq \frac{1}{\sigma - 1},$$

and the falsity of the theorem would contradict Theorem 7–18.

By now the proof of the following analog of Theorem 7–10 should be a simple exercise for the reader.

THEOREM 7–21.   *For $\sigma > 1$,*

$$-3 \frac{L'}{L} (\sigma, \chi_0) - 4 \operatorname{Re} \frac{L'}{L} (\sigma + ti, \chi) - \operatorname{Re} \frac{}{L} (\sigma + 2ti, \chi^2) \geq 0.$$

Theorem 7–13 becomes

THEOREM 7–22.   *There is a $c_1(k) \geq 8$ such that $L(s, \chi) \neq 0$ for $t > c_1(k)$ and $\sigma \geq 1 - c_2/\log t$.*

The only difference in the proofs is that now the first inequality of Theorem 7–8 is applied with $f(s) = L(s, \chi^2)$ and $s_0 = \sigma + 2ti$, while the second is applied with $f(s) = L(s, \chi)$ and $s_0 = \sigma + ti$. A' the constants now depend on k. After these trivial mod ..ions, the proofs are identical.

Similarly, replacing $\zeta(s)$ by $L(s, \chi)$ throughout, Theorem 7–14 becomes

THEOREM 7-23. *For $t \geq c_9(k) > 8$ and $\sigma \geq 1 - c_8 (\log t)^{-1}$,*
$|\log L(s, \chi)| < \log^2 t$.

The constant $c_9(k)$ may be different from the c_9 of Theorem 7-14; the subscript is retained to facilitate reference to Fig. 7-3. In the same way, c_{11} becomes $c_{11}(k)$.

Instead of proceeding directly to the analog of Theorem 7-15, it is convenient to break the argument into two steps.

THEOREM 7-24. *For $(k, l) = 1$, we have*

$$\sum_{\substack{p \leq x \\ p \equiv l \,(\mathrm{mod}\ k)}} \log \frac{x}{p} = \frac{1}{2\pi i h} \sum_{\chi} \frac{1}{\chi(l)} \int_{(2)} \frac{x^s}{s^2} \log L(s, \chi)\, ds + O(\sqrt{x} \log^2 x).$$

Proof: Using Theorem 7-9 and the series expansion for $\log L(s, \chi)$, we obtain

$$\frac{1}{2\pi i} \int_{(2)} \frac{x^s}{s^2} \log L(s, \chi)\, ds = \frac{1}{2\pi i} \int_{(2)} \frac{x^s}{s^2} \sum_{m,p} \frac{\chi(p^m)}{mp^{ms}}\, ds$$

$$= \frac{1}{2\pi i} \sum_{m,p} \frac{\chi(p^m)}{m} \int_{(2)} \frac{(x/p^m)^s}{s^2}\, ds$$

$$= \sum_{\substack{m,p \\ p^m \leq x}} \frac{\chi(p^m) \log (x/p^m)}{m}$$

$$= \sum_{\substack{p \leq x \\ p \nmid k}} \chi(p) \log \frac{x}{p} + \sum_{\substack{p^m \leq x \\ m \geq 2}} \frac{\chi(p^m) \log (x/p^m)}{m}$$

$$= \sum_{\substack{p \leq x \\ p \nmid k}} \chi(p) \log \frac{x}{p} + O(\sqrt{x} \log^2 x).$$

Multiplying by $1/\chi(l)$ and summing over all characters modulo k, we deduce with the help of Theorem 6-7 that

$$\sum_{\chi} \frac{1}{\chi(l)} \sum_{\substack{p \leq x \\ p \nmid k}} \chi(p) \log \frac{x}{p} = h \sum_{\substack{p \leq x \\ p \equiv l \,(\mathrm{mod}\ k)}} \log \frac{x}{p}$$

$$= \frac{1}{2\pi i} \sum_{\chi} \frac{1}{\chi(l)} \int_{(2)} \frac{x^s}{s^2} \log L(s, \chi)\, ds + O(\sqrt{x} \log^2 x),$$

which is the theorem. (Here and throughout the remainder of this section, the implied constant in the O-symbol may depend on k.)

To estimate the integrals appearing in Theorem 7–24, we must distinguish two cases. First consider the case $\chi = \chi_0$. Every property of the integrand which was used in estimating

$$\int_{(2)} \frac{x^s}{s^2} \log \zeta(s)\, ds$$

carries over to the integrand of

$$\int_{(2)} \frac{x^s}{s^2} \log L(s, \chi_0)\, ds.$$

It follows that for suitable c with $0 < c < 1$,

$$\int_{(2)} \frac{x^s}{s^2} \log L(s, \chi_0)\, ds = 2\pi i \int_c^1 \frac{x^s}{s^2}\, ds + O(xe^{-\alpha\sqrt{\log x}}).$$

On the other hand, if $\chi \neq \chi_0$, then $L(s, \chi)$ has no pole at $s = 1$, but the other properties used earlier still obtain. Hence, if we do not cut the plane, but consider the line segments Γ_5 and $\bar{\Gamma}_5$ in Fig. 7–3 as a single segment Γ_8, and omit Γ_6, Γ_7, and $\bar{\Gamma}_6$, then the function

$$\psi(s, \chi) = \frac{x^s}{s^2} \log L(s, \chi)$$

is regular in the region bounded by Γ_1, Γ_2, Γ_3, Γ_4, Γ_8, $\bar{\Gamma}_4$, $\bar{\Gamma}_3$, $\bar{\Gamma}_2$, $\bar{\Gamma}_1$, so that

$$\int_{(2)} \psi(s, \chi)\, ds = \left(\int_{2-\infty i}^{2-ui} - \int_{\bar{\Gamma}_2 + \bar{\Gamma}_3 + \bar{\Gamma}_4 + \Gamma_8 + \Gamma_4 + \Gamma_3 + \Gamma_2} + \int_{2+ui}^{2+\infty i} \right) \psi(s, \chi)\, ds.$$

Moreover, the integral along each of these new arcs either tends to zero or is

$$O(xe^{-\alpha\sqrt{\log x}}).$$

It follows that

$$\sum_{\substack{p \le x \\ p \equiv l \,(\mathrm{mod}\, k)}} \log \frac{x}{p} = \frac{1}{h} \int_c^1 \frac{x^s\, ds}{s^2} + O(xe^{-\alpha\sqrt{\log x}}), \tag{16}$$

which is the analog of Theorem 7–15. In exactly the same way as Theorem 7–16 was deduced from Theorem 7–15, equation (16) leads to the desired result:

THEOREM 7–25. *If k is a fixed integer and $(k, l) = 1$, then, as $x \to \infty$,*

$$\pi(x; k, l) = \frac{1}{\varphi(k)} \int_2^x \frac{du}{\log u} + O(xe^{-\frac{1}{2}\alpha\sqrt{\log x}}).$$

As consequences of Theorem 7–25, we have that

$$\pi(x; k, l) \sim \frac{1}{\varphi(k)} \frac{x}{\log x},$$

and that, if $(k, l_1) = (k, l_2) = 1$, then

$$\lim_{x \to \infty} \frac{\pi(x; k, l_1)}{\pi(x; k, l_2)} = 1,$$

so that asymptotically there are equally many primes in the progressions $kt + l_1$ and $kt + l_2$.

A serious drawback of Theorem 7–25 is that the error term is not uniform in k. This precludes applying this version of the theorem to problems in which k increases with x, and these unfortunately are among the most important applications of this kind of theorem. It is known that the error term in Theorem 7–25 is uniform in k for $k < \log^m x$ for some $m > 0$, in other words, that the relation displayed in the theorem can be used if k increases sufficiently slowly with x. The proof of the more general theorem, while similar to that given here, is more complicated. The chief difficulty is this: when dealing with fixed k, it is enough to prove that $L(s, \chi) \neq 0$ for $s = 1 + ti$ in order to deduce that for some $c_{11}(k)$, $L(s, \chi) \neq 0$ for $1 - c_{11} < \sigma \leq 1$, $|t| < c_9(k)$. When k increases, however, c_{11} might tend to zero quite rapidly as a function of k, in which case the integral along Γ_8 would not be negligible. It is therefore necessary to investigate further the zeros of the L-functions near the line $\sigma = 1$ for small $|t|$.

7–5 The integers representable as a sum of two squares. As a final illustration of the methods of this chapter, we shall obtain an asymptotic estimate for $B(x)$, the number of integers not exceeding x which can be written as a sum of two squares. The integers counted are exactly those in whose prime-power factorization the primes $r \equiv 3 \pmod 4$ occur only to even powers.* The following heuristic argument indicates that it is to be expected that $B(x)$ is of the order of magnitude of $x/\sqrt{\log x}$, which is in agreement with the result to be obtained.

Take x very large. Since one out of every p integers is divisible by p, the number of integers up to x not divisible by p is about

---

* Cf. Volume I, Theorem 7–3.

$x(1 - 1/p)$. Hence the number not divisible by any $p \leq \sqrt{x}$ is roughly

$$x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right),$$

so that, by the Prime Number Theorem,

$$x \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right) \approx \frac{x}{\log x},$$

where the symbol "$\approx$" means "is probably of the order of magnitude of." To count the integers contributing to $B(x)$, we do not want to eliminate all composite numbers, but only those divisible by an odd power of any of the various primes $r \equiv 3 \pmod 4$. As in the cross-classification principle,* we can omit all those divisible by r, then reintroduce those divisible by r^2, then take out those divisible by r^3, etc., giving

$$x \left(1 - \frac{1}{r}\right)\left(1 + \frac{1}{r^2}\right)\left(1 - \frac{1}{r^3}\right) \cdots$$

as the number left after accounting for the one prime r. (The product has only finitely many factors.) Hence

$$B(x) \approx x \prod_{r \leq \sqrt{x}} \left(1 - \frac{1}{r}\right) \prod_{r^2 \leq \sqrt{x}} \left(1 + \frac{1}{r^2}\right) \cdots,$$

and since each product after the first converges as $x \to \infty$, we can write simply

$$B(x) \approx x \prod_{r \leq \sqrt{x}} \left(1 - \frac{1}{r}\right).$$

Now

$$\log \prod_{p \leq \sqrt{x}} \left(1 - \frac{1}{p}\right) = \sum_{p \leq \sqrt{x}} \log \left(1 - \frac{1}{p}\right) \approx \sigma \log x,$$

and since, by the results of the preceding section, about half the p's are r's, we have

$$\log \prod_{r \leq \sqrt{x}} \left(1 - \frac{1}{r}\right) \approx -\frac{1}{2} \log \log x = -\log \sqrt{\log x},$$

---

* Cf. Volume I, Theorem 6-4.

so that
$$B(x) \approx \frac{x}{\sqrt{\log x}} \cdot$$

Probably the most that can be said for this argument is that after seeing it, the reader should not be very surprised to learn that, for some $b > 0$,

$$B(x) \sim \frac{bx}{\sqrt{\log x}} \cdot \qquad (17)$$

Nevertheless, it is just this type of reasoning which underlies the proof of (17) which will now be developed.

If we put

$$b_n = \begin{cases} 1 & \text{if } n = x^2 + y^2 \text{ for some integers } x, y, \\ 0 & \text{otherwise,} \end{cases}$$

then
$$B(x) = \sum_{n \le x} b_n.$$

For $\sigma > 1$ let

$$f(s) = \sum_{n=1}^{\infty} \frac{b_n}{n^s};$$

the series converges absolutely in this domain, and uniformly in any closed bounded region to the right of the line $\sigma = 1$. Using q and r to denote primes congruent to 1 and 3 (mod 4) respectively, we deduce from the definition of b_n and Theorem 6–3 that, for $\sigma > 1$,

$$f(s) = \left(1 + \frac{1}{2^s} + \frac{1}{2^{2s}} + \cdots\right) \prod_q \left(1 + \frac{1}{q^s} + \frac{1}{q^{2s}} + \cdots\right)$$

$$\times \prod_r \left(1 + \frac{1}{r^{2s}} + \frac{1}{r^{4s}} + \cdots\right)$$

$$= (1 - 2^{-s})^{-1} \prod_q (1 - q^{-s})^{-1} \prod_r (1 - r^{-2s})^{-1}.$$

As was pointed out in formula (5) of the preceding chapter,

$$\zeta(s)L(s) = (1 - 2^{-s})^{-1} \prod_q (1 - q^{-s})^{-2} \prod_r (1 - r^{-2s})^{-1}, \qquad (18)$$

where $L(s) = L(s, \chi)$ is the L-function for the nonprincipal character (mod 4),

$$\chi(n) = \begin{cases} 0 & \text{if } 2 \mid n, \\ (-1)^{\frac{1}{2}(n-1)} & \text{if } 2 \nmid n. \end{cases}$$

(The relation (18) was proved earlier only for $s > 1$ and real; the extension to the half-plane $\sigma > 1$ is immediate.) Hence, for $\sigma > 1$,

$$f^2(s) = (1 - 2^{-s})^{-1} \prod_r (1 - r^{-2s})^{-1} \zeta(s) L(s). \qquad (19)$$

Since L is regular for $\sigma > 0$ and

$$L(1) = 1 - \frac{1}{3} + \frac{1}{5} - \cdots = \frac{\pi}{4},$$

the function ζL has a simple pole at $s = 1$, with residue $\pi/4$, but is otherwise regular for $\sigma > 0$. Moreover, neither $\zeta(s)$ nor $L(s)$ is zero for s in the region Q of Fig. 7–3, for suitable positive c_{11} and c_9. Since the functions

$$(1 - 2^{-s})^{-1} \qquad \text{and} \qquad \prod_r (1 - r^{-2s})^{-1} \qquad (20)$$

are regular and different from zero for $\sigma > \frac{1}{2}$, and bounded in absolute value for $\sigma \geq \sigma_0 > \frac{1}{2}$, we deduce the following properties of f from known properties of ζ and L.

THEOREM 7–26.   (a) $f^2(s)$ *is regular and different from zero in the region Q of Fig. 7–3, for suitable c_{11} and c_9, and it has a simple pole at $s = 1$, with residue*

$$\frac{\pi}{2} \prod_r (1 - r^{-2})^{-1} = b^2. \qquad (21)$$

Hence f is also regular in Q, and $f^2(s) \cdot (s - 1)$ is regular in the uncut region Q' formed from Q by omitting Γ_6, Γ_7, and $\bar{\Gamma}_6$, and joining Γ_5 and $\bar{\Gamma}_5$.
    (b) *For $|t| \geq 8$ and s in Q, the inequality $|f(s)| < c_{14} \log |t|$ holds* (cf. Theorem 7–11 (b)).

From this follows the usual consequence.

THEOREM 7–27.   *For suitable $c < 1$,*

$$\sum_{n \leq x} b_n \log \frac{x}{n} = \frac{1}{\pi i} \int_c^1 \frac{x^s}{s^2} f(s) \, ds + O(x e^{-\alpha \sqrt{\log x}}).$$

Proof: The proof follows the lines of that of Theorem 7–15 as regards changing the path of integration in the relation

$$\sum_{n \leq x} b_n \log \frac{x}{n} = \frac{1}{2\pi i} \int_{(2)} \frac{x^s}{s^2} f(s) \, ds$$

and estimating the new integrals along those paths which are

bounded away from $s = 1$; the only change is that the estimate $|f(s)| < c_{14} \log |t|$ is used rather than $|\log \zeta(s)| < \log^2 |t|$. Omitting the tiresome details, we arrive at the relation

$$\sum_{n=1}^{x} b_n \log \frac{x}{n} = \frac{1}{2\pi i} \left(-\int_{\bar{\Gamma}_6} - \int_{\Gamma_7} - \int_{\Gamma_6} \right) \frac{x^s}{s^2} f(s) \, ds + O(xe^{-\alpha\sqrt{\log x}}).$$

In the neighborhood of $s = 1$, $f(s)$ has the expansion

$$f(s) = \frac{b}{\sqrt{s-1}} + \cdots,$$

with b as in (21). Here $\sqrt{s-1} > 0$ for $s > 1$. Putting $s = 1 + \delta e^{i\theta}$, we have

$$\int_{\Gamma_7} \frac{x^s}{s^2} f(s) \, ds = O\left(\frac{x^{1+\delta}}{(1-\delta)^2} \cdot \frac{1}{\sqrt{\delta}} \cdot 2\pi\delta \right) = o(1)$$

as $\delta \to 0$.

Since $f^2(s)(s-1)$ is single-valued in Q', the quantity $2 \arg f(s) + \arg (s-1)$ is unchanged by traversing a path in Q from $\bar{\Gamma}_6$ to Γ_6. Since $\arg (s-1)$ increases by 2π, $\arg f(s)$ decreases by π, so that $f(s)$ has opposite signs on the two edges of the cut. Hence

$$\int_{\Gamma_6} \frac{x^s}{s^2} f(s) \, ds + \int_{\bar{\Gamma}_6} \frac{x^s}{s^2} f(s) \, ds = 2 \int_{\bar{\Gamma}_6} \frac{x^s}{s^2} f(s) \, ds,$$

and

$$\sum_{n=1}^{x} b_n \log \frac{x}{n} = \frac{1}{\pi i} \int_{1-c_{11}}^{1} \frac{x^s}{s^2} f(s) \, ds + O(xe^{-\alpha\sqrt{\log x}}).$$

The proof is complete.

THEOREM 7-28. *As* $x \to \infty$,

$$B(x) = \frac{Bx}{\sqrt{\log x}} + O\left(\frac{x}{(\log x)^{\frac{3}{4}}} \right),$$

where

$$B = \frac{1}{\sqrt{2}} \prod_r \left(1 - \frac{1}{r^2} \right)^{-\frac{1}{2}}.$$

Proof: On $\bar{\Gamma}_6$ we have

$$\frac{f(s)}{s^2} = \frac{bi}{\sqrt{1-s}\,(1-(1-s))^2} + O(\sqrt{1-s})$$

$$= \frac{bi}{\sqrt{1-s}} + O(\sqrt{1-s})$$

as $s \to 1^-$, so that

$$\sum_{n=1}^{x} b_n \log \frac{x}{n}$$

$$= \frac{1}{\pi i} \int_{1-c_{11}}^{1} \frac{x^s b i}{\sqrt{1-s}} \, ds + O\left(\int_{1-c_{11}}^{1} x^s \sqrt{1-s} \, ds\right) + O\left(\frac{x}{\log^2 x}\right)$$

$$= \frac{b}{\pi} \int_{0}^{c_{11}} x^{1-u} u^{-\frac{1}{2}} \, du + O\left(\int_{0}^{c_{11}} x^{1-u} u^{\frac{1}{2}} \, du\right) + O\left(\frac{x}{\log^2 x}\right)$$

$$= \frac{bx}{\pi} \int_{0}^{c_{11}} e^{-u \log x} u^{-\frac{1}{2}} \, du + O\left(x \int_{0}^{c_{11}} e^{-u \log x} u^{\frac{1}{2}} \, du\right) + O\left(\frac{x}{\log^2 x}\right)$$

$$= \frac{bx}{\pi} \int_{0}^{c_{11} \log x} e^{-v} \left(\frac{v}{\log x}\right)^{-\frac{1}{2}} \frac{dv}{\log x} + O\left(x \int_{0}^{c_{11} \log x} e^{-v} \left(\frac{v}{\log x}\right)^{\frac{1}{2}} \frac{dv}{\log x}\right)$$

$$+ O\left(\frac{x}{\log^2 x}\right)$$

$$= \frac{bx}{\pi \sqrt{\log x}} \int_{0}^{c_{11} \log x} e^{-v} v^{-\frac{1}{2}} \, dv + O\left(\frac{x}{\log^{\frac{3}{2}} x} \int_{0}^{\infty} e^{-v} v^{\frac{1}{2}} \, dv\right) + O\left(\frac{x}{\log^2 x}\right)$$

$$= \frac{bx}{\pi \sqrt{\log x}} \left(\Gamma\left(\frac{1}{2}\right) - \int_{c_{11} \log x}^{\infty} e^{-v} v^{-\frac{1}{2}} \, dv\right) + O\left(\frac{x}{\log^{\frac{3}{2}} x}\right)$$

$$= \frac{bx}{\pi \sqrt{\log x}} \left\{\sqrt{\pi} + O\left(\int_{c_{11} \log x}^{\infty} e^{-v} \, dv\right)\right\} + O\left(\frac{x}{\log^{\frac{3}{2}} x}\right).$$

Thus

$$\sum_{n=1}^{x} b_n \log \frac{x}{n} = \frac{Bx}{\sqrt{\log x}} + O\left(\frac{x}{\log^{\frac{3}{2}} x}\right),$$

where

$$B = \frac{b}{\sqrt{\pi}} = \frac{1}{\sqrt{2}} \prod_{r} (1 - r^2)^{-\frac{1}{2}}.$$

Now let $\delta = \delta(x)$ be positive. Then

$$\sum_{n=1}^{x+\delta x} b_n \log \frac{x + \delta x}{n} - \sum_{n=1}^{x} b_n \log \frac{x}{n}$$

$$= \log (1 + \delta) \sum_{n=1}^{x} b_n + \sum_{n=x}^{x+\delta x} b_n \log \frac{x + \delta x}{n}$$

$$= \log (1 + \delta) B(x) + O\left(\log (1 + \delta) \cdot \delta x\right),$$

while

$$\frac{Bx\,(1+\delta)}{\sqrt{\log\,(x+\delta x)}} - \frac{Bx}{\sqrt{\log\,x}} = \frac{Bx}{\sqrt{\log\,x}}\left(\frac{1+\delta}{\sqrt{\log\,(x+\delta x)/\log\,x}} - 1\right)$$

$$= \frac{Bx}{\sqrt{\log\,x}}\left(\frac{1+\delta}{\sqrt{1+\log\,(1+\delta)/\log\,x}} - 1\right)$$

$$= \frac{Bx}{\sqrt{\log\,x}}\left(\frac{1+\delta}{1+O(\delta/\log\,x)} - 1\right)$$

$$= \frac{Bx}{\sqrt{\log\,x}}\left(\frac{\delta + O(\delta/\log\,x)}{1+O(\delta/\log\,x)}\right)$$

$$= \frac{Bx}{\sqrt{\log\,x}}\,(\delta + O(\delta/\log\,x)).$$

$$+ \delta) = \delta + O(\delta^2) \text{ as } \delta \to 0,$$

$$B(x) = \frac{}{\sqrt{\log\,x}}\left\{\frac{\delta}{\log\,(1+\delta)} + O\left(\frac{1}{\log\,x}\right)\right\} + O(\delta x) + O\left(\frac{x}{\delta\,\log^{\frac{3}{2}}x}\right)$$

$$= \frac{Bx}{\sqrt{\log\,x}}\,(1 + O(\delta)) + O\left(\frac{x}{\log^{\frac{3}{2}}x}\right) + O(\delta x) + O\left(\frac{x}{\delta\,\log^{\frac{3}{2}}x}\right).$$

Choosing $\delta(x) = \log^{-\frac{3}{4}}x$, we obtain

$$B(x) = \frac{Bx}{\sqrt{\log\,x}} + O\left(\frac{x}{\log^{\frac{5}{4}}x}\right) + O\left(\frac{x}{\log^{\frac{3}{2}}x}\right)$$

$$+ O\left(\frac{x}{\log^{\frac{3}{4}}x}\right) + O\left(\frac{x}{\log^{\frac{3}{4}}x}\right)$$

$$= \frac{Bx}{\sqrt{\log\,x}} + O\left(\frac{x}{\log^{\frac{3}{4}}x}\right),$$

and the proof is complete.

SUPPLEMENTARY READING

Chapter 1

DICKSON, L. E., *Introduction to the Theory of Numbers*, Chicago: University of Chicago Press, 1929.

DICKSON, L. E., *Modern Elementary Theory of Numbers*, Chicago, University of Chicago Press, 1939.

FORD, L. R., *An Introduction to the Theory of Automorphic Functions*, London: George Bell & Sons, Ltd., 1915. Reprinted, Chelsea Publishing Company, New York, 1951.

JONES, B. W., *The Arithmetic Theory of Quadratic Forms*, Carus Mathematical Monograph #10, Buffalo, N.Y.: Mathematical Association of America, 1950. Distributed by John Wiley & Sons, Inc., New York.

KLEIN, F., *Vorlesungen über die Theorie der Elliptischen Modulfunktionen*, Leipzig: Teubner Verlagsgesellschaft, 1890-1892.

Chapter 2

HECKE, E., *Vorlesungen über die Theorie der Algebraischen Zahlen*, Leipzig: Akademische Verlagsgesellschaft m.b.H., 1923. Reprinted, Chelsea Publishing Company, New York, 1948.

LANDAU, E., *Vorlesungen über Zahlentheorie*, vol. 3, Leipzig: S. Hirzel Verlag, 1927. Reprinted, Chelsea Publishing Company, New York, 1947.

POLLARD, H., *The Theory of Algebraic Numbers*, Carus Mathematical Monograph #9, Buffalo, N.Y.: Mathematical Association of America, 1950. Distributed by John Wiley & Sons, Inc., New York.

REID, L. W., *Elements of the Theory of Algebraic Numbers*, New York: The Macmillan Company, 1910.

WEYL, H., *Algebraic Theory of Numbers*, Annals of Mathematics Studies, #1, Princeton: Princeton University Press, 1940.

Chapter 3

LANDAU, E., *Vorlesungen über Zahlentheorie*, vol. 3.

MORDELL, L. J., *Three Lectures on Fermat's Last Theorem*, New York: Cambridge University Press, 1921.

Chapter 5

GELFOND, A. O., *The Approximation of Algebraic Numbers by Algebraic Numbers and the Theory of Transcendental Numbers*, American Mathematical Society Translation #65, Providence: American Mathematical Society, 1952. Translated from *Uspekhi Matematicheskikh Nauk* (Moscow) **4**, no. 4 (32), 19–49 (1949).

KOKSMA, J. F., *Diophantische Approximationen*, Berlin: Springer-Verlag OHG, 1936. (Ergebnisse der Mathematik, vol. 4, no. 4.) Reprinted, Chelsea Publishing Company, New York, 1951.

PERRON, O., *Irrationalzahlen*, 2nd edition, Berlin: Walter De Gruyter & Co., 1929.

SIEGEL, C. L., *Transcendental Numbers*, Annals of Mathematics Studies, #16, Princeton: Princeton University Press, 1949.

Chapter 6

HASSE, H., *Vorlesungen über Zahlentheorie*, Berlin: Springer-Verlag OHG, 1950.

LANDAU, E., *Vorlesungen über Zahlentheorie*, vol. 1.

Chapter 7

ESTERMANN, T., *Introduction to Modern Prime Number Theory*, Cambridge Tracts, #41, New York: Cambridge University Press, 1952.

INGHAM, A. E., *The Distribution of Prime Numbers*, Cambridge Tracts, #30, New York: Cambridge University Press, 1932.

LANDAU, E., *Vorlesungen über Zahlentheorie*, vol. 2.

LANDAU, E., *Handbuch der Lehre von der Verteilung der Primzahlen*, Leipzig: Teubner Verlagsgesellschaft, 1909. Reprinted, Chelsea Publishing Company, New York, 1953.

LANDAU, E., *Einführung in die Elementare und Analytische Theorie der Algebraischen Zahlen und der Ideale*, Leipzig: Teubner Verlagsgesellschaft, 1918. Reprinted, Chelsea Publishing Company, New York, 1949.

LIST OF SYMBOLS

INDEX